

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 4, 2005, 17:57:59 : Search time 8169 Seconds  
(without alignments)  
11504.580 Million cell updates/sec

Title: US-09-508-658B-1

Perfect score: 2036

Sequence: 1 agaccggggagagcggcgsgg.....aaaatacaaaaattagctg 2036

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database :

1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sca:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2027	99.6	2027	9	AB006682 Homo sapi
2	2027	99.6	2258	6	CQ731864 Sequence 1
3	1969	96.7	2245	6	A97284 Sequence 1
4	1969	96.7	2245	6	A97284 Sequence 1
5	1025	50.3	1545	6	CQ731863 Sequence
6	1025	50.3	1545	6	AB006683 Homo sapi
7	692	34.0	1463	9	AB006685 Homo sapi
8	336	16.5	2586	9	HS1TPK X80853 H. sapiens g
9	336	16.5	20000	9	AB006684 Homo sapi
10	336	16.5	42133	9	AP001060 Homo sapi
11	336	16.5	340000	9	AP001754 Homo sapi
12	322	15.8	36284	9	HS19610 Homo sapi
13	269	13.2	110000	2	AC003656_5 Continuation (6 of
14	195	9.6	143192	2	AL954247 Pan trogl
15	60	2.9	60	6	CQ541728 Sequence
16	50	2.5	101	11	BV184386 sqm14455
17	32	1.6	558	11	BV163805 RPAMMSQ0
18	32	1.6	574	11	BV089350 RPAMMSQ0
19	32	1.6	593	11	BV088992 RPAMMSQ0

20	32	1.6	1656	10	AF128772 Mus muscu
21	32	1.6	1656	10	AF128773 Mus muscu
22	32	1.6	1659	6	A97293 Sequence 10
23	32	1.6	1659	10	MMU122243
24	32	1.6	1744	10	AF128121 Mus muscu
25	32	1.6	1747	10	AF128120 Mus muscu
26	32	1.6	1756	10	AF128119 Mus muscu
27	32	1.6	1759	10	AF128118 Mus muscu
28	32	1.6	1906	10	MMU243821
29	32	1.6	1921	10	AF128117 Mus muscu
30	32	1.6	1924	10	AF128116 Mus muscu
31	32	1.6	1933	10	AF128115 Mus muscu
32	32	1.6	1936	10	AF079536 Mus muscu
33	32	1.6	18351	10	AF105002 Mus muscu
34	32	1.6	18616	10	MMU007715
35	32	1.6	46872	10	AF073797 Mus muscu
36	32	1.6	158049	2	AC015891 Mus muscu
37	32	1.5	496	11	BV088997 RPAMMSQ0
38	30	1.5	1622	10	AF128125 Mus muscu
39	30	1.5	1625	10	AF128124 Mus muscu
40	30	1.5	1634	10	AF128123 Mus muscu
41	30	1.5	1637	10	AF128122 Mus muscu
42	30	1.5	125105	10	AC007433 Mus muscu
43	30	1.5	217688	2	AC108592
44	27	1.3	193047	9	AC120780 Pan trogl
45	27	1.3	194191	2	AC097332 Pan trogl
46	27	1.3	206315	2	AC148550 Callithr
47	27	1.3	212879	2	AC148552 Callithr
48	26	1.3	577	11	BV089370 RPAMMSQ0
49	26	1.3	3932	6	C0843504 Sequence
50	26	1.3	3932	6	AK126488 Homo sapi
51	26	1.3	38715	9	AC000092 Homo sapi
52	26	1.3	49087	2	AF293359 Homo sapi
53	26	1.3	65608	2	AC132197 Homo sapi
54	26	1.3	67706	2	AC073517 Homo sapi
55	26	1.3	69586	2	AC135170 Homo sapi
56	26	1.3	73421	9	AL391378 Human DNA
57	26	1.3	99682	9	HS452817
58	26	1.3	105731	9	AC105430 Homo sapi
59	26	1.3	109149	2	AP000681 Homo sapi
60	26	1.3	109867	2	AC004898 Homo sapi
61	26	1.3	110000	2	AL583823_0
62	26	1.3	112752	2	AF030876 Homo sapi
63	26	1.3	114991	9	AC010331 Homo sapi
64	26	1.3	126787	9	AC008102 Homo sapi
65	26	1.3	128726	9	AL136458 Human DNA
66	26	1.3	131915	2	AC090170 Homo sapi
67	26	1.3	131975	9	AC008940 Homo sapi
68	26	1.3	142895	9	AC012435 Homo sapi
69	26	1.3	144117	9	AC092574 Homo sapi
70	26	1.3	154702	2	AC011471 Homo sapi
71	26	1.3	154906	2	AC026254 Homo sapi
72	26	1.3	155216	2	AC141275 Homo sapi
73	26	1.3	155216	2	AC141313 Homo sapi
74	26	1.3	156491	9	AL591720 Human DNA
75	26	1.3	156911	9	AC009335 Homo sapi
76	26	1.3	159693	2	AC009781 Homo sapi
77	26	1.3	160584	2	AC022248 Homo sapi
78	26	1.3	163025	2	AP000758 Homo sapi
79	26	1.3	163652	2	AC145845 Pan trogl
80	26	1.3	165106	2	AC016034 Homo sapi
81	26	1.3	165447	9	AC027575 Homo sapi
82	26	1.3	167190	9	AC023301 Homo sapi
83	26	1.3	169904	2	AC132072 Papio anu
84	26	1.3	170404	2	AP002441 Homo sapi

C 93	26	1.3	170940	2	AC025132	166	AC025132 Homo sapi
C 94	26	1.3	171480	9	AC008937	167	AC008937 Homo sapi
C 95	26	1.3	171998	9	AC026120	168	AC026120 Homo sapi
C 96	26	1.3	172192	9	AC091067	169	AC091067 Homo sapi
C 97	26	1.3	172357	2	AP001808	170	AP001808 Homo sapi
C 98	26	1.3	173735	2	AC027063	171	AC027063 Homo sapi
C 99	26	1.3	173987	9	AL4392084	172	AL4392084 Human DNA
C 100	26	1.3	173987	9	AL450337	173	AL450337 Human DNA
C 101	26	1.3	174739	9	AL356428	174	AL356428 Human DNA
C 102	26	1.3	177483	9	AC108046	175	AC108046 Homo sapi
C 103	26	1.3	178058	9	AC009139	176	AC009139 Homo sapi
C 104	26	1.3	178384	2	AL591026	177	AL591026 Homo sapi
C 105	26	1.3	180754	2	AC148963	178	AC148963 Papi'o anu
C 106	26	1.3	181445	9	AC026367	179	AC026367 Homo sapi
C 107	26	1.3	181552	9	AC012417	180	AC012417 Homo sapi
C 108	26	1.3	182258	9	AC025212	181	AC025212 Homo sapi
C 109	26	1.3	183293	2	AC148883	182	AC148883 Pan trogl
C 110	26	1.3	183549	9	AL445473	183	AL445473 Human DNA
C 111	26	1.3	185882	2	AC073554	184	AC073554 Homo sapi
C 112	26	1.3	187099	2	AC018673	185	AC018673 Homo sapi
C 113	26	1.3	192078	2	AP001649	186	AP001649 Homo sapi
C 114	26	1.3	192429	2	AC149185	187	AC149185 Papi'o anu
C 115	26	1.3	192515	2	AC023497	188	AC023497 Homo sapi
C 116	26	1.3	194241	2	AC093324	189	AC093324 Homo sapi
C 117	26	1.3	197079	2	AC147385	190	AC147385 Pan trogl
C 118	26	1.3	198146	2	AL590422	191	AL590422 Homo sapi
C 119	26	1.3	198312	9	AL590422	192	AL590422 Human DNA
C 120	26	1.3	199558	2	AP001807	193	AP001807 Homo sapi
C 121	26	1.3	200771	2	AC093368	194	AC093368 Homo sapi
C 122	26	1.3	202490	2	AP002426	195	AP002426 Homo sapi
C 123	26	1.3	205466	9	AC010548	196	AC010548 Homo sapi
C 124	26	1.3	207520	9	AC091138	197	AC091138 Homo sapi
C 125	26	1.3	208726	9	AP006287	198	AP006287 Homo sapi
C 126	26	1.3	211047	2	AP001461	199	AP001461 Homo sapi
C 127	26	1.3	225957	9	AC069175	200	AC069175 Homo sapi
C 128	26	1.3	283951	2	AC109383		AC109383 Rattus no
C 129	25	1.2	51	11	BV190786		BV190786 sqmnl6842
C 130	25	1.2	51	11	BV190815		BV190815 sqmnl6849
C 131	25	1.2	51	11	BV190816		BV190816 sqmnl6849
C 132	25	1.2	172	6	AX900402		AX900402 Sequence
C 133	25	1.2	172	6	BD055935		BD055935 Sequence
C 134	25	1.2	179	9	AF205206		AF205206 Homo sapi
C 135	25	1.2	225	11	G33040		G33040 A010A10 Hum
C 136	25	1.2	230	6	CO691192		CO691192 Sequence
C 137	25	1.2	239	6	CO694874		CO694874 Sequence
C 138	25	1.2	239	6	AX908330		AX908330 Sequence
C 139	25	1.2	239	6	BD043853		BD043853 Sequence
C 140	25	1.2	268	6	BD244152		BD244152 Determina
C 141	25	1.2	282	6	CO458030		CO458030 Sequence
C 142	25	1.2	284	9	HS4343113		HS4343113 Homo sapi
C 143	25	1.2	292	6	BD060313		BD060313 Secretd
C 144	25	1.2	305	9	AF680019		AF680019 Homo sapi
C 145	25	1.2	311	9	HUMCG1A1B		MI0626 Human AluI
C 146	25	1.2	323	6	CO664850		CO664850 Sequence
C 147	25	1.2	347	6	CO669831		CO669831 Sequence
C 148	25	1.2	357	6	AX070928		AX070928 Sequence
C 149	25	1.2	377	11	BV188826		BV188826 sqmnl6136
C 150	25	1.2	401	6	AX270430		AX270430 Sequence
C 151	25	1.2	401	6	AX271961		AX271961 Sequence
C 152	25	1.2	401	11	BV194224		BV194224 sqmnl8029
C 153	25	1.2	401	11	BV194725		BV194725 sqmnl8210
C 154	25	1.2	403	6	CO525656		CO525656 Sequence
C 155	25	1.2	403	6	AY270591		AY270591 Homo sapi
C 156	25	1.2	413	6	AR443222		AR443222 Sequence
C 157	25	1.2	413	6	AY270421		AY270421 Homo sapi
C 158	25	1.2	419	6	CO694882		CO694882 Sequence
C 159	25	1.2	427	6	CO485396		CO485396 Sequence
C 160	25	1.2	437	11	BV191711		BV191711 sqmnl7173
C 161	25	1.2	447	6	AX386886		AX386886 Sequence
C 162	25	1.2	447	6	AX875349		AX875349 Sequence
C 163	25	1.2	447	6	BD155411		BD155411 Primer fo
C 164	25	1.2	467	6	CO515219		CO515219 Sequence
C 165	25	1.2	472	6	CO463932		CO463932 Sequence

166	25	1.2	474	6	CO679548	166	CO679548 Sequence
167	25	1.2	486	6	AR423153	167	AR423153 Sequence
168	25	1.2	486	6	AX983847	168	AX983847 Sequence
169	25	1.2	486	6	BD118706	169	BD118706 EST and e
170	25	1.2	489	11	G27575	170	G27575 human STS S
C 171	25	1.2	496	6	AX875149		AX875149 Sequence
C 172	25	1.2	496	6	BD155211		BD155211 Primer fo
C 173	25	1.2	497	6	AR416239		AR416239 Sequence
C 174	25	1.2	497	6	AX973073		AX973073 Sequence
C 175	25	1.2	497	6	BD111792		BD111792 EST and e
C 176	25	1.2	520	6	CO423219		CO423219 Sequence
C 177	25	1.2	530	6	AR414297		AR414297 Sequence
C 178	25	1.2	530	6	AX971131		AX971131 Sequence
C 179	25	1.2	530	6	BD109850		BD109850 EST and e
C 180	25	1.2	541	6	AX136754		AX136754 Sequence
C 181	25	1.2	541	6	BD123994		BD123994 Secretory
C 182	25	1.2	550	6	AX385831		AX385831 Sequence
C 183	25	1.2	555	6	AX872371		AX872371 Sequence
C 184	25	1.2	555	6	AX872526		AX872526 Sequence
C 185	25	1.2	555	6	BD152433		BD152433 Primer fo
C 186	25	1.2	582	6	BD152588		BD152588 Primer fo
C 187	25	1.2	582	6	CO519529		CO519529 Sequence
C 188	25	1.2	590	6	AX385157		AX385157 Sequence
C 189	25	1.2	601	11	BV184089		BV184089 sqmnl4299
C 190	25	1.2	601	11	BV187598		BV187598 sqmnl5479
C 191	25	1.2	601	11	BV197012		BV197012 sqmnl9130
C 192	25	1.2	605	6	CO526165		CO526165 Sequence
C 193	25	1.2	619	9	AR327081		AR327081 Homo sapi
C 194	25	1.2	678	9	HS4342777		HS4342777 Homo sapi
C 195	25	1.2	684	6	CO409544		CO409544 Sequence
C 196	25	1.2	685	6	AR202069		AR202069 Sequence
C 197	25	1.2	700	6	AX182860		AX182860 Sequence
C 198	25	1.2	700	6	AX182861		AX182861 Sequence
C 199	25	1.2	703	6	CO510762		CO510762 Sequence
C 200	25	1.2	716	9	AY620649		AY620649 Saguinus

## ALIGNMENTS

RESULT 1  
LOCUS AB006682  
DEFINITION Homo sapiens APECD mRNA for AIRE-1, complete cds.  
ACCESSION AB006682.1 GI:2696614  
VERSION  
KEYWORDS APECD, AIRE-1.  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Mammalia; Eutheria; Chordata; Carnivora; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Nagamine,K., Peterson,P., Scott,H.S., Kudoh,J., Minoshima,S.,  
Heino,M., Kichu,K.U.E., Lalitoti,M.D., Mullis,P.E.,  
Antonarakis,S.E., Kawasaki,K., Asakawa,S., Ito,F. and Shimizu,N.  
Positional cloning of the APECD gene

TITLE  
JOURNAL Nat. Genet. 17 (4), 393-398 (1997)  
MEDLINE 98061086  
PUBMED 9398839

REFERENCE  
AUTHORS 2 (bases 1 to 2027)  
TITLE Direct Submission  
JOURNAL Submitted (16-AUG-1997) Nobuyoshi Shimizu, Keio University School  
of Medicine, Department of Molecular Biology, 35 Shinanomachi,  
Shinjuku-ku, Tokyo 160, Japan (E-mail:shimizu@med.keio.ac.jp,  
Tel:03-3351-2370, Fax:03-3351-2370)

FEATURES  
source location/Qualifiers  
1..2027  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/isolate="Caucasian"  
/db\_xref="taxon:9606"  
/chromosome="21"



QY 1750 TCCGCGGCGCCCTTCCCTCCTGAGCCCGAGATGCGGGAGCATGAGCTGTAGAG 1809  
DB 1741 TCCGCGGCGCCCTTCCCTCCTGAGCCCGAGATGCGGGAGCATGAGCTGTAGAG 1800  
QY 1810 AGTGTGAGAGAGACATCTCTTCTCTGAGTCTGGAAGCGCGCTGGATCAAGAG 1869  
DB 1801 AGTGTGAGAGAGACATCTCTTCTCTGAGTCTGGAAGCGCGCTGGATCAAGAG 1860  
QY 1870 GGGAGCGCGCACCTCTGTGAGTGTGCTGCTGTAACAGCTGTGTTCTGGGGACAC 1929  
DB 1861 GGGAGCGCGCACCTCTGTGAGTGTGCTGCTGTAACAGCTGTGTTCTGGGGACAC 1920  
QY 1930 CAGCCATCATGTGCTGGAATTAACCTGCCCCCACTTCTCTACTGGAAGTCCCGG 1989  
DB 1921 CAGCCATCATGTGCTGGAATTAACCTGCCCCCACTTCTCTACTGGAAGTCCCGG 1980  
QY 1990 GAGCTCTCTCTGCTGCTGCTGAGCTACTAAATAATATAATTAAGCTG 2036  
DB 1981 GAGCTCTCTCTGCTGCTGCTGAGCTACTAAATAATATAATTAAGCTG 2027

RESULT 2  
CQ731864 2258 bp DNA linear PAT 03-FEB-2004  
LOCUS CQ731864  
DEFINITION Sequence 17798 from Patent WO02068579.  
ACCESSION CQ731864  
VERSION CQ731864.1 GI:42309571  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof  
JOURNAL Patent: WO 02068579-A 17798 06-SEP-2002;  
PE Corporation (US)

FEATURES  
source Location/Qualifiers  
1..2258

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 99.6%; Score 2027; DB 6; Length 2258;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AGACGGGCGGGCGCAGAGCGCGGAGGCGCCCAAGCCCGCGGAGCCGAGGCCAA 69  
DB 1 AGACGGGCGGGCGCAGAGCGCGGAGGCGCCCAAGCCCGCGGAGCCGAGGCCAA 60  
QY 70 GCGAGGGGCTGCGAGTGTCCCGGAGCCACCGGCTCCGCCAGCCCGGGTCCCGCGC 129  
DB 61 GCGAGGGGCTGCGAGTGTCCCGGAGCCACCGGCTCCGCCAGCCCGGGTCCCGCGC 120  
QY 130 CCAACCCATGAGCGAGCGCGGCGCTACCGCGCTTCTGAGGCTGACCGCAGGAGAT 189  
DB 121 CCAACCCATGAGCGAGCGCGGCGCTACCGCGCTTCTGAGGCTGACCGCAGGAGAT 180  
QY 190 CGGCGTGGCGGTGAGAGAGCGCCCTTCCACCTGCTGCAAGCGCTGAGCAACGACTGAT 249  
DB 181 CGGCGTGGCGGTGAGAGAGCGCCCTTCCACCTGCTGCAAGCGCTGAGCAACGACTGAT 240  
QY 250 CCCCGAGGACAAGTTTCAAGAGAGCGTTCTATCTGAAGAAAAGAGAGGCTCCCGAGGC 309  
DB 241 CCCCGAGGACAAGTTTCAAGAGAGCGTTCTATCTGAAGAAAAGAGAGGCTCCCGAGGC 300  
QY 310 CTTTCAAGCGCCCTCTGTCTGTGCTGCTGAGTCCAGGACTTCCAGCGCATCTTCTTG 369

DB 301 CTTTCAAGCGCCCTCTGTCTGTGCTGAGTCCAGGAGCTTCCAGCGCATCTTCTTG 360  
QY 370 GAGGGTGTGTTCAAGAGACTACACTGAGAGGCTATGGCGGGCTGAGGCCATCTTGA 429  
DB 361 GAGGGTGTGTTCAAGAGACTACACTGAGAGGCTATGGCGGGCTGAGGCCATCTTGA 420  
QY 430 CAGCTTCCCAAGATGTGAGACTTCAAGCCAGCCCGGAGAGGGAGAGAGCCCGGCGT 489  
DB 421 CAGCTTCCCAAGATGTGAGACTTCAAGCCAGCCCGGAGAGGGAGAGAGCCCGGCGT 480  
QY 490 CCCCAAGGCTTGTGATCGGCACCCAGACTCCCAACCAAGAGAGAGGCTCAAGAGAGC 549  
DB 481 CCCCAAGGCTTGTGATCGGCACCCAGACTCCCAACCAAGAGAGAGGCTCAAGAGAGC 540  
QY 550 TCGAGTGGCGGCGCAGAGCCCTGACTCCAAAGGGGACCCGACGCCAGGCTCAACT 609  
DB 541 TCGAGTGGCGGCGCAGAGCCCTGACTCCAAAGGGGACCCGACGCCAGGCTCAACT 600  
QY 610 GAAGGCGAAGCCCGCAAGAAAGCCGAGAGCAGCCAGAGCAGCGCTTCACTCGG 669  
DB 601 GAAGGCGAAGCCCGCAAGAAAGCCGAGAGCAGCCAGAGCAGCGCTTCACTCGG 660  
QY 670 GAACGGATTCAGACCATGTCTGAGTTCAGTCCAGAGAGCTGTGCTCCGGGGA 729  
DB 661 GAACGGATTCAGACCATGTCTGAGTTCAGTCCAGAGAGCTGTGCTCCGGGGA 720  
QY 730 CGTCCCGGAGCGCAGAGGGGCGGTGAGGGAGATCTCATCCAGAGGTGTTGAGTCAAG 789  
DB 721 CGTCCCGGAGCGCAGAGGGGCGGTGAGGGAGATCTCATCCAGAGGTGTTGAGTCAAG 780  
QY 790 CGGCTTCAAGAGTGCATTCAGGTTGCGGGAGTTCTACACTCCAGCAAGTTGAGAGA 849  
DB 781 CGGCTTCAAGAGTGCATTCAGGTTGCGGGAGTTCTACACTCCAGCAAGTTGAGAGA 840  
QY 850 CTCGCGCACTGGGAGAGAACAGAGCCCGCAGCAGCAGTGGCCCGGAAGCTCTGTGAGC 909  
DB 841 CTCGCGCACTGGGAGAGAACAGAGCCCGCAGCAGCAGTGGCCCGGAAGCTCTGTGAGC 900  
QY 910 CAAGGAGAGCCAGGGGCGCTGCGCCCGGTGAGAGTGAAGCTGAGCGGCGCAGAGGGCAG 969  
DB 901 CAAGGAGAGCCAGGGGCGCTGCGCCCGGTGAGAGTGAAGCTGAGCGGCGCAGAGGGCAG 960  
QY 970 CGTTCCCGCCCTCTGCGCCCTTCCCAAGTACCCCGAGCTTCCACGAAGAAATGAGAGCA 1029  
DB 961 CGTTCCCGCCCTCTGCGCCCTTCCCAAGTACCCCGAGCTTCCACGAAGAAATGAGAGCA 1020  
QY 1030 GTGTGCGGTGTGTGGGAGAGGCGGGAGCTCATCTGTGTGACGCGTCCCTCGGCGCTT 1089  
DB 1021 GTGTGCGGTGTGTGGGAGAGGCGGGAGCTCATCTGTGTGACGCGTCCCTCGGCGCTT 1080  
QY 1090 CCACGTGGCTGCTGTGCTCCCTCGCTCCGGAGAGATCCCAAGTGGGACTTGAAGGTGCTC 1149  
DB 1081 CCACGTGGCTGCTGTGCTCCCTCGCTCCGGAGAGATCCCAAGTGGGACTTGAAGGTGCTC 1140  
QY 1150 CAGCTGCTGAGGAGCAAGTTCAGAGAGTGTGAGCCCGCGGAGAGAGCCCGGCCCA 1209  
DB 1141 CAGCTGCTGAGGAGCAAGTTCAGAGAGTGTGAGCCCGCGGAGAGAGCCCGGCCCA 1200  
QY 1210 GGAAGCACCCGTGTGAGAGATCCCGCTTCCCGCGGGCTTGTGTGCGGGAGAGAGTAA 1269  
DB 1201 GGAAGCACCCGTGTGAGAGATCCCGCTTCCCGCGGGCTTGTGTGCGGGAGAGAGTAA 1260  
QY 1270 AGGTCAACTGGGGAAACCTTACCGCGGAGTGAAGACGACTCTGTCTAAGACACTGCG 1329  
DB 1261 AGGTCAACTGGGGAAACCTTACCGCGGAGTGAAGACGACTCTGTCTAAGACACTGCG 1320  
QY 1330 GGTCCGCTTCTGTGAGCCCGCTGACAGGGCTGTGACTCTTGTGCGCTGACCCCTACT 1389  
DB 1321 GGTCCGCTTCTGTGAGCCCGCTGACAGGGCTGTGACTCTTGTGCGCTGACCCCTACT 1380  
QY 1390 GTGTGTGGTCTTGAAGGTCAAGAGACTTGTGCTTGTGTGCGGCTGTGGGCTGTGCGG 1449  
DB 1381 GTGTGTGGTCTTGAAGGTCAAGAGACTTGTGCTTGTGTGCGGCTGTGGGCTGTGCGG 1440



OY	1450	AGATGGTAAGGACGCGTGGCCGCTGTACTCATTGCGCGCGCTCCCTTCCACATGCGCCTGGCA	1509
Db	1441	AGATGGTAAGGACGCGTGGCCGCTGTACTCATTGCGCGCGCTCCCTTCCACATGCGCCTGGCA	1500
OY	1510	CTTCCACGCGGACCTCTCCGGCCCGGAGCGGGCCTGGCTGCAGATCTCTCTCAGAGAGA	1569
Db	1501	CTTCCACGCGGACCTCTCCGGCCCGGAGCGGGCCTGGCTGCAGATCTCTCTCAGAGAGA	1560
OY	1570	CGTGAACCCAGACCCCTGTGTGAGGGGGTGTGTGCCCCCAGCCCCCGCGCTGGCCCCCTGG	1629
Db	1561	CGTGAACCCACCCCTGTGTGAGGGGGTGTGTGCCCCCAGCCCCCGCGCTGGCCCCCTGG	1620
OY	1630	GCCTGGCAAGATGACATGTGCACGTGACAGACCCCGCTTGCACAGGGATGACCTGGAGTC	1689
Db	1621	GCCTGGCAAGATGACATGTGCACGTGACAGACCCCGCTTGCACAGGGATGACCTGGAGTC	1680
OY	1690	CCTTCTGAGCAGACACACTTGCATGGCATCTGTGAGTGGGCAATCCAGACATGGACCG	1749
Db	1681	CCTTCTGAGCAGACACACTTGCATGGCATCTGTGAGTGGGCAATCCAGACATGGACCG	1740
OY	1750	TCCGGCGGCCCCCTTCCCCCTCTCTGACCCCAATGGCCGGGACATGCAGCTCTGATGAGAG	1809
Db	1741	TCCGGCGGCCCCCTTCCCCCTCTCTGACCCCAATGGCCGGGACATGCAGCTCTGATGAGAG	1800
OY	1810	AGTGTGAGGAAGACACCTCTCTCTGATCTCTGGAAGCCGGCGGCTGGGATGCAAGAG	1869
Db	1801	AGTGTGAGGAAGACACCTCTCTCTGATCTCTGGAAGCCGGCGGCTGGGATGCAAGAG	1860
OY	1870	GGGACAGCGCACCTTGTGATGAGTCTGGCGCTGTAAACAGCTCTGTGTTTCTGGGGACAC	1929
Db	1861	GGGACAGCGCACCTTGTGATGAGTCTGGCGCTGTAAACAGCTCTGTGTTTCTGGGGACAC	1920
OY	1930	CAGCCATCATGTGCTGTGAAATTAACCCCTGCCCACTTCTCTACTGTGAAATGCCCG	1989
Db	1921	CAGCCATCATGTGCTGTGAAATTAACCCCTGCCCACTTCTCTACTGTGAAATGCCCG	1980
OY	1990	GAGGCTCTCCCTGTGCTGTGATCTTAATAAATTAATAAATTAACCTG	2036
Db	1981	GAGGCTCTCCCTGTGCTGTGATCTTAATAAATTAATAAATTAACCTG	2027

RESULT 3				
LOCUS	A97284			
DEFINITION	A97284	2245 bp	DNA	linear
ACCESSION	A97284	Sequence 1 from Patent WO9918197.		PAT 26-JAN-2000
VERSION	A97284.1	GI:6780668		
KEYWORDS				
SOURCE	unidentified			
ORGANISM	unidentified			
REFERENCE	unclassified.			
AUTHORS	1 (bases 1 to 2245)			
TITLE	Yaspo, M. and Lehnach, H.			
	NUCLEIC ACID MOLECULE ENCODING A (POLY)PEPTIDE CO-SEGREGATING IN			
	MUTATED FORM WITH AUTOIMMUNE POLYENDOCRINOPATHY CANDIDIASIS			
	ECUODERMAL DYSTROPHY (APECED)			
	Patent: WO 9918197-A 1 15-APR-1999;			
JOURNAL	MAX PLANCK GESELLSCHAFT (DE); YASPO MARIE LAURE (DE)			
FEATURES	location/Qualifiers			
source	1..2245			
	/organism="unidentified"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32644"			

Query	Match	Similarity	Score	1969;	DB	6;	Length	2245;
	Best	Local	Similarity	100.04;	Pred.	No.	0;	
	Matches	2019;	Conservative	0;	Mismatches	1;	Indels	0;
								Gaps
Qy	17	CGGGCGCACACCGGGCGGAGGCCCCACACCCCCCGCGGAGCCCGAGGCCAAGCGAGGG	76					
Qb	1	CGGGCGCACACCGGGCGGAGGCCCCACACCCCCCGCGGAGCCCGAGGCCAAGCGAGGG	60					

QY	77	GCTGCGACATGTCCTCCGGGAGCCACCGCGTCGCGCCAGCGCCCGAGTCTCCGCGCCACCCC	136
Db	61	GCTGCGACATGTCCTCCGGGAGCCACCGCGTCGCGCCAGCGCCCGAGTCTCCGCGCCACCCC	120
QY	137	ATGGCGACGGACGCGGCGCTACGCGCGGCTTCTGAGGCTGACACGCGACGAGATTCGCGGTG	196
Db	121	ATGGCGACGGACGCGGCGCTACGCGCGGCTTCTGAGGCTGACACGCGACGAGATTCGCGGTG	180
QY	197	GCCGTGACAGCGGCTTCCCACTGCTGCAACGCGCTGCGCTGACACAGACGTGGTCCCCGAG	256
Db	181	GCCGTGACAGCGGCTTCCCACTGCTGCAACGCGCTGCGCTGACACAGACGTGGTCCCCGAG	240
QY	257	GACAAAGTTTCAGAGACGCTTCATCTGAAAGAAAAGAGGGCTGCGCCCCAGGCTTCCAC	316
Db	241	GACAAAGTTTCAGAGACGCTTTCATCTGAAAGAAAAGAGGGCTGCGCCCCAGGCTTCCAC	300
QY	317	GCCCTCTGTCTCTGCGCTGCTAGCCAGAGATTCACAGCCATCTGAGACTTCTGAGAGGTG	376
Db	301	GCCCTCTGTCTCTGCGCTGCTAGCCAGAGATTCACAGCCATCTGAGACTTCTGAGAGGTG	360
QY	377	CTGTTCAAGGACTAACACTGAGGGGCTATGCGCGGCTGACACCATCCTTGACAGGCTTC	436
Db	361	CTGTTCAAGGACTAACACTGAGGGGCTATGCGCGGCTGACACCATCCTTGACAGGCTTC	420
QY	437	CCCAAGATGTGACCTGACGCGCGCCGGAAGGGAGAGACCCCGCGCTGCCCAAG	496
Db	421	CCCAAGATGTGACCTGACGCGCGCCGGAAGGGAGAGACCCCGCGCTGCCCAAG	480
QY	497	GCTTTGGTATCCGCGCACCCAGACTCCCCACACAGAGAGAGGCTCAAGAGAGGCTGAGCT	556
Db	481	GCTTTGGTATCCGCGCACCCAGACTCCCCACACAGAGAGAGGCTCAAGAGAGGCTGAGCT	540
QY	557	GCGGCGCAGCAGCGCTGACTCCAGAGGGGACCGCGCAGCGCTCTGAACTGAAGGCG	616
Db	541	GCGGCGCAGCAGCGCTGACTCCAGAGGGGACCGCGCAGCGCTCTGAACTGAAGGCG	600
QY	617	AAGCCCCCAAGAAAGCCGGAAGACGCGACGCGACAGCAGCGCTTTCACCTCGGAAACGG	676
Db	601	AAGCCCCCAAGAAAGCCGGAAGACGCGACGCGACAGCAGCGCTTTCACCTCGGAAACGG	660
QY	677	ATTGAGACATATGTCAGCTTCACTCCAGAGAGCTGTGGCAATCTCTCCGCGGGAACGTCGCG	736
Db	661	ATTGAGACATATGTCAGCTTCACTCCAGAGAGCTGTGGCAATCTCTCCGCGGGAACGTCGCG	720
QY	737	GAGCGCCAGAGGGCGCTGAGAGGGAGTCCCTCATCCAGCAGGTGTTTGAAGTCAAGCGGCTCC	796
Db	721	GAGCGCCAGAGGGCGCTGAGAGGGAGTCCCTCATCCAGCAGGTGTTTGAAGTCAAGCGGCTCC	780
QY	797	AAGAAATGCAATCCAGGTTGGCGGGGAGTTCTACACTCCAGCAAGTTGGAAGACTTCGCGC	856
Db	781	AAGAAATGCAATCCAGGTTGGCGGGGAGTTCTACACTCCAGCAAGTTGGAAGACTTCGCGC	840
QY	857	AGTGGGAAGAACAAAGGCGCGGACGACGAGTGGCCGGAAGGCTCTGATTCGAGCCAAAGGA	916
Db	841	AGTGGGAAGAACAAAGGCGCGGACGACGAGTGGCCGGAAGGCTCTGATTCGAGCCAAAGGA	900
QY	917	GCCCAAGGCGCTGCCCTCCGATGGAAGGTAGGCTTGAGCGCAGCAGGCGACGCTTCC	976
Db	901	GCCCAAGGCGCTGCCCTCCGATGGAAGGTAGGCTTGAGCGCAGCAGGCGACGCTTCC	960
QY	977	GCCCTCTGTGCGCTTCCCAATGACCCCCAGACTCCACAGAAATGAGAGACAGATGTGCG	1033
Db	961	GCCCTCTGTGCGCTTCCCAATGACCCCCAGACTCCACAGAAATGAGAGACAGATGTGCG	1021
QY	1037	GTTGATCGGAGACGCGCGGAGACTCATCTGCTGTGACGCGCTGCGCTCGGAGCTTCCACCTG	1099
Db	1021	GTTGATCGGAGACGCGCGGAGACTCATCTGCTGTGACGCGCTGCGCTCGGAGCTTCCACCTG	1088
QY	1097	GCTGACCTGTCTCCCTTCGCTCCGCGGAGATTCGCCAGTGGAGACTGAGAGTGTCTCAAGTGC	1155
Db	1081	GCTGACCTGTCTCCCTTCGCTCCGCGGAGATTCGCCAGTGGAGACTGAGAGTGTCTCAAGTGC	1144

Accession	Sequence	Position
27		1080
28		1081
29		1082
30		1083
31		1084
32		1085
33		1086
34		1087
35		1088
36		1089
37		1090
38		1091
39		1092
40		1093
41		1094
42		1095
43		1096
44		1097
45		1098
46		1099
47		1100
48		1101
49		1102
50		1103
51		1104
52		1105
53		1106
54		1107
55		1108
56		1109
57		1110
58		1111
59		1112
60		1113
61		1114
62		1115
63		1116
64		1117
65		1118
66		1119
67		1120
68		1121
69		1122
70		1123
71		1124
72		1125
73		1126
74		1127
75		1128
76		1129
77		1130
78		1131
79		1132
80		1133
81		1134
82		1135
83		1136
84		1137
85		1138
86		1139
87		1140
88		1141
89		1142
90		1143
91		1144
92		1145
93		1146
94		1147
95		1148
96		1149
97		1150
98		1151
99		1152
100		1153
101		1154
102		1155
103		1156
104		1157
105		1158
106		1159
107		1160
108		1161
109		1162
110		1163
111		1164
112		1165
113		1166
114		1167
115		1168
116		1169
117		1170
118		1171
119		1172
120		1173
121		1174
122		1175
123		1176
124		1177
125		1178
126		1179
127		1180
128		1181
129		1182
130		1183
131		1184
132		1185
133		1186
134		1187
135		1188
136		1189
137		1190
138		1191
139		1192
140		1193
141		1194
142		1195
143		1196
144		1197
145		1198
146		1199
147		1200
148		

QY 1157 CTGACGAAACAGTCCAGGAGGTGAGCCCGGGCAGAGAGCCCGGCCCCAGAGCCA 1216  
DB 1141 CTGACGAAACAGTCCAGGAGGTGAGCCCGGGCAGAGAGCCCGGCCCCAGAGCCA 1200  
QY 1217 CCCGTGAGAACCCCGCTCCCGGGGCTTAAGTCCGCGGAGAGAGAGGTAAGAGTCCA 1276  
DB 1201 CCCGTGAGAACCCCGCTCCCGGGGCTTAAGTCCGCGGAGAGAGAGGTAAGAGTCCA 1260  
QY 1277 CCGTGGGAAACCCCTAGCCGAGCATGACAGACTCTTGTCTACAGAGACCTGCCGCTCCG 1336  
DB 1261 CCGTGGGAAACCCCTAGCCGAGCATGACAGACTCTTGTCTACAGAGACCTGCCGCTCCG 1320  
QY 1337 CTTTCTGACGCCCGCTGCGAGGCTGAGACTCTTCTGCGCCCTGACGCCCTTACTGTGTGTG 1396  
DB 1321 CTTTCTGACGCCCGCTGCGAGGCTGAGACTCTTCTGCGCCCTGACGCCCTTACTGTGTGTG 1380  
QY 1397 GGTCTTGAAGGTTCAGAGAAACCTGGGTCTGTGGTCCCGGCTTGGCGGGGTGTGCGAGATGGT 1456  
DB 1381 GGTCTTGAAGGTTCAGAGAAACCTGGGTCTGTGGTCCCGGCTTGGCGGGGTGTGCGAGATGGT 1440  
QY 1457 ACGGAGTGTGCGGTGTAATCACTGCGCGCTGCGCTTCCACTGCGGCTGCCACTTCCCA 1516  
DB 1441 ACGGAGTGTGCGGTGTAATCACTGCGCGCTGCGCTTCCACTGCGGCTGCCACTTCCCA 1500  
QY 1517 GCGCGACCTCCCGGCGCGGAGCGGCTGCGCTGAGATCTGTCTCAGAGAGAGCTGAGCC 1576  
DB 1501 GCGCGACCTCCCGGCGCGGAGCGGCTGCGCTGAGATCTGTCTCAGAGAGAGCTGAGCC 1560  
QY 1577 CCAAGCCCTGTGAGAGGGGTGTGCTGGGCCCCCAGCCCGGCTTGGCCCTTGGGCTTGGC 1636  
DB 1561 CCAAGCCCTGTGAGAGGGGTGTGCTGGGCCCCCAGCCCGGCTTGGCCCTTGGGCTTGGC 1620  
QY 1637 AAGGATGACATGCGCAGTACGAGAGCCGCTCTGACAGAGGATGACTGAGTCTTCTG 1696  
DB 1621 AAGGATGACATGCGCAGTACGAGAGCCGCTCTGACAGAGGATGACTGAGTCTTCTG 1680  
QY 1697 ACGGAGCACACTTGCATGAGCATCTGAGTGGGCGCATCCAGAGATGAGCCGCTGCGGCG 1756  
DB 1681 ACGGAGCACACTTGCATGAGCATCTGAGTGGGCGCATCCAGAGATGAGCCGCTGCGGCG 1740  
QY 1757 GCGCCCTTCCCTCCGACCCGAGATGGCGCGGACATGCACTCTTATGAGAGAGTGTG 1816  
DB 1741 GCGCCCTTCCCTCCGACCCGAGATGGCGCGGACATGCACTCTTATGAGAGAGTGTG 1800  
QY 1817 AGAAGGACACTCTCTCTCACTGAGTGGGAGCGGCGGCTGGGATCAAGAAAGGAGAG 1876  
DB 1801 AGAAGGACACTCTCTCTCTCACTGAGTGGGAGCGGCGGCTGGGATCAAGAAAGGAGAG 1860  
QY 1877 CGGCACCTTGTGAGTGTGCTGGGTGTAACAGCTGTGTCTGAGGAGACACAGCCAT 1936  
DB 1861 CGGCACCTTGTGAGTGTGCTGGGTGTAACAGCTGTGTCTGAGGAGACACAGCCAT 1920  
QY 1937 CATGTGCGTGGAAATTAATTAACCTGCGGCACTTCTTACTCTGGAATCCCCCGGAGACTTC 1996  
DB 1921 CATGTGCGTGGAAATTAATTAACCTGCGGCACTTCTTACTCTGGAATCCCCCGGAGACTTC 1980  
QY 1997 TCGTTCCTGTGAGTCACTTAATAATTAATTAATTAAGTCTG 2036  
DB 1981 TCGTTCCTGTGAGTCACTTAATAATTAATTAATTAAGTCTG 2020

RESULT 4  
HSAPECED 2245 bp mRNA linear PRI 18-JUN-2000  
LOCUS Homo Sapiens mRNA for AIRE protein.  
ACCESSION Z97990.1 GI:2665370  
VERSION 297990.1  
KEYWORDS AIRE protein.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1

AUTHORS  
Aaltonen,J., Bjrees,P., Perheentupa,J., Horelli-Kuitunen,N.,  
Palotci,A., Peltonen,L., Lee,Y.S., Francis,F., Hennig,S., Thiel,C.,  
Lehach,H. and Yaspo,M.L.  
TITLE  
An autoimmune disease, APECED, caused by mutations in a novel gene  
featuring two PHD-type zinc finger domains  
JOURNAL  
Nat. Genet. 17, 399-403 (1997)  
MEDLINE  
98061087  
REFERENCE  
2 (bases 1 to 2245)  
AUTHORS  
Yaspo,M.L.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (21-JUN-1997) Max Planck Institut fur Molekulare Genetik,  
Imnestraase 73, Berlin D-14195, Germany  
REMARK  
revised by submitter 24-SEP-1997  
FEATURES  
location/Qualifiers  
source  
1..2245  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/tissue\_type="thymus"  
1..1810  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="B1-1"  
/tissue\_type="thymus"  
/clone\_lib="lambda GT11 from Clontech (cat# h15010B)"  
1811..2180  
/clone\_lib="lambda GT11 from Clontech (cat# h15010B)"  
121..1758  
/codon\_start=1  
/product="AIRE protein"  
/protein\_id="CA10790.1"  
/db\_xref="GI:2665371"  
/db\_xref="GOA:O43918"  
/db\_xref="Swiss-Prot:O43918"  
/translation="MATDALRLRLHRTETAVNDSAFPULHAIADHVPEDKFO  
ETLHKEKGGPOAFAHLSMLTQDSTALDFMVLPRDYLVBYGRLOPLDSFK  
DVLDSOPRKRPAPVPAKALVPPRLPTKRASEARAAPALPRGASPSQLKA  
KPPKRPSSABEQRLPLNGIOTMSASQRAVMSGVPARGAVEGILIOQVESG  
GSKCTIQVGEFTYTSKFEDSGSKNKAASSGPKPLVRAKAGAGAGGAGARLQO  
GSVPALALPSDPQJHKNEDCAVCRDGBELICDGGPRAFHACLSPPLEIPTS  
WKSCLQATVQVQRAEPRPQEPVETPLPQLRAGAEVRGPPEPLAGMTTL  
VYKHLPAFSAAPPLGLDSALHPLLCVGPBEOQLAGARGVCGDGDVLRCTHCA  
AAFWHCHPAGTSPRGTLRCRSGSDVTPAPVAGVLAPPARLAPPAODTASHE  
PALHDDLESLSLSEHTPDGILQMALQSMARAPPPS"

ORIGIN  
Query Match 96.7%; Score 1969; DB 9; Length 2245;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 17 CGGCGCACAGCCGCGCGAGGCCCCACAGCCCGCGGAGCCGAGGCGCAAGGAGG 76  
DB 1 CGGCGCACAGCCGCGCGAGGCCCCACAGCCCGCGGAGCCGAGGCGCAAGGAGG 60  
QY 77 GCTGCAGTGTCCCGGAGCCCAAGCGGTCCGCGAGCCCGGAGTCCCGGCGCAAGG 136  
DB 61 GCTGCAGTGTCCCGGAGCCCAAGCGGTCCGCGAGCCCGGAGTCCCGGCGCAAGG 120  
QY 137 ATGCGACGGAAGCGCGGCTGACCGGCTTCTGAGGCTGACCGGACGAGAGATCGG 196  
DB 121 ATGCGACGGAAGCGCGGCTGACCGGCTTCTGAGGCTGACCGGACGAGAGATCGG 180  
QY 197 GCGTGGACAGCGGCTTCCCACTGTGTGACAGCGGCTGAGTACGAGAGTGTG 256  
DB 197 GCGTGGACAGCGGCTTCCCACTGTGTGACAGCGGCTGAGTACGAGAGTGTG 256

[illegible]

QY	1337	CGTTCTGAGGCGCGCTGCGCAGGGGCTGGACTCTTCGCGCCCTGACACCCCTTACTGTGTG	1396
Db	1321	CGTTCTGAGGCGCGCTGCGCAGGGGCTGGACTCTTCGCGCCCTGACACCCCTTACTGTGTG	1380
QY	1397	GGTCCTGAGGGTCAGCAAACTTGCTCTTGATGCGCGTTGCGGGGTGTGTCGAGATGCT	1456
Db	1381	GGTCCTGAGGGTCAGCAAACTTGCTCTTGATGCGCGTTGCGGGGTGTGTCGAGATGCT	1440
QY	1457	ACGACGCGTGGCGGGGTGAACCTCATGCGCGCTGCTCCACTGAGCGCTGCACTTCCCA	1516
Db	1441	ACGACGCGTGGCGGGGTGAACCTCATGCGCGCTGCTCCACTGAGCGCTGCACTTCCCA	1500
QY	1517	GCCGCGACCTTCGCGCCCGGGACGCGGCTGCGCTGACAGATCTCTCTCAGAGAGCGTGACC	1576
Db	1501	GCCGCGACCTTCGCGCCCGGGACGCGGCTGCGCTGACAGATCTCTCTCAGAGAGCGTGACC	1560
QY	1577	CCACCCCCCTGTGAGAGGGGGGTGTGCGCCCGACGCCCGCCCGGCGGCGCTGCGGCTGCGC	1636
Db	1561	CCACCCCCCTGTGAGAGGGGGGTGTGCGCCCGACGCCCGCCCGGCGGCGCTGCGGCTGCGC	1620
QY	1637	AAGATGACACTGCGCAGTGCAGACGAGCCCGCTCTGACAGAGATGACTGTGAGTCCCTTCTG	1696
Db	1621	AAGATGACACTGCGCAGTGCAGACGAGCCCGCTCTGACAGAGATGACTGTGAGTCCCTTCTG	1680
QY	1697	AGCGAGCAACCTTTCGATGGCATCTCTGCAAGTGGGCCATCCAGACATGCGCCGTCGCGCG	1756
Db	1681	AGCGAGCAACCTTTCGATGGCATCTCTGCAAGTGGGCCATCCAGACATGCGCCGTCGCGCG	1740
QY	1757	GCCCCCTTCCGCTCTGACCCGACATGAGCGGGACATGACAGCTGTGATGAGAGAGTCTG	1816
Db	1741	GCCCCCTTCCGCTCTGACCCGACATGAGCGGGACATGACAGCTGTGATGAGAGAGTCTG	1800
QY	1817	AGAGAGCAACCTCTTCTCTCACTCTTGAAAGCCGCGCGGCTGCGATCAAGAGGGGACAG	1876
Db	1801	AGAGAGCAACCTCTTCTCTCACTCTTGAAAGCCGCGCGGATCAAGAGGGGACAG	1860
QY	1877	CGCAACCTTGTGAGTGCCTCGGCTGTAACAGCTGTGTGTTTGTGGGAGCACAGGCAT	1936
Db	1861	CGCAACCTTGTGAGTGCCTCGGCTGTAACAGCTGTGTGTTTGTGGGAGCACAGGCAT	1920
QY	1937	CATGTGCTGTGAATTAACCCCTGCGCCCACTTCTCTACTCTGGAAGTCCCGGGAGGCTC	1996
Db	1921	CATGTGCTGTGAATTAACCCCTGCGCCCACTTCTCTACTCTGGAAGTCCCGGGAGGCTC	1980
QY	1997	TGCTTGGCTGTGACTTAATAAATTAATAAATTAACCTG 2036	
Db	1981	TGCTTGGCTGTGACTTAATAAATTAATAAATTAACCTG 2020	

RESULT 5				
CQ731863				
LOCUS	CQ731863	1545 bp	DNA	linear
DEFINITION	Sequence 17797 from Patent WO02068579.			
ACCESSION	CQ731863			
VERSION	CQ731863.1	GI:42309567		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Cetrarchii; Homnidae; Homo.			
REFERENCE				
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.			
TITLE	Rits, such as nucleic acid arrays, comprising a majority of			
	humanexons or transcripts, for detecting expression and other uses			
	thereof			
JOURNAL	Patent: WO 02068579-A 17797 06-SEP-2002;			
	PE Corporation (NY) (US)			
FEATURES				
source	location/Qualifiers			
	1..1545			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			



```

Db      581 CGGCTGCGCTCGGCGCTTCACCTGGCGTGGCTCCGCTCCGCGCTCGGAGATCCCGAC 640
Qy      1132 TGGGACCTGGAGGTGCTCCAGCTGCTGACGACACAGTCCAGAGAGTGCAGCCCGGCG 1191
Db      641 TGGGACCTGGAGGTGCTCCAGCTGCTGACGACACAGTCCAGAGAGTGCAGCCCGGCG 700
Qy      1192 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1251
Db      701 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
Qy      1252 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1311
Db      761 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 820
Qy      1312 TGTCTAACAAGACCTGCGCGCTTCGCTGCAAGCGCGCGCTGCGCGCTGCACTGCGCGCTGC 1371
Db      821 TGTCTAACAAGACCTGCGCGCTTCGCTGCAAGCGCGCGCTGCGCGCTGCACTGCGCGCTGC 880
Qy      1372 GCGCGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
Db      881 GCGCGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
Qy      1432 GCGTTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1491
Db      941 GCGTTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1000
Qy      1492 CTTTCACTGGGCGCTGCACTTCCAGCGCGGACCTTCCGCGCGCGGAGAGAGAGAGAGAGAG 1551
Db      1001 CTTTCACTGGGCGCTGCACTTCCAGCGCGGACCTTCCGCGCGCGGAGAGAGAGAGAGAGAG 1060
Qy      1552 CAGATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1611
Db      1061 CAGATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1120
Qy      1612 CGCGCGCGCTGCGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1671
Db      1121 CGCGCGCGCTGCGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1180
Qy      1672 CAGGAGATGACTGAGAGTCCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1731
Db      1181 CAGGAGATGACTGAGAGTCCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1240
Qy      1732 CATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1791
Db      1241 CATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1300
Qy      1792 ATGCAGCTCTGATGAGAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1851
Db      1301 ATGCAGCTCTGATGAGAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1360
Qy      1852 CGGCGCTGGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1911
Db      1361 CGGCGCTGGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1420
Qy      1912 CTGCTGTTCTGGGAGACAGGAGCATGATGCTGGAATTTAAACCTGCGCGAGAGAGAGAGAGAG 1971
Db      1421 CTGCTGTTCTGGGAGACAGGAGCATGATGCTGGAATTTAAACCTGCGCGAGAGAGAGAGAGAG 1480
Qy      1972 TACTCTGGAAGTCCCGGAGAGCTCTCTTGGCTGGTGAAGCTTAAATAATAATAATAATAAT 2031
Db      1481 TACTCTGGAAGTCCCGGAGAGCTCTCTTGGCTGGTGAAGCTTAAATAATAATAATAATAAT 1540
Qy      2032 AGCTG 2036
Db      1541 AGCTG 1545

```

```

RESULT 7
AB006685 1463 bp mRNA linear PRI 13-DEC-1997
LOCUS     AB006685 Homo sapiens APECD mRNA for AIR-3, complete cds.
DEFINITION
ACCESSION AB006685
VERSION    AB006685.1 GI:2696622

```

```

KEYWORDS  APECD; AIR-3.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (sites)
            Nagamine,K., Peterson,P., Scott,H.S., Kudoh,J., Minoshima,S.,
            Helmo,M., Krohn,K.J.E., Lalioi,M.D., Mullis,P.E.,
            Antonarakis,S.E., Kawaaki,K., Asakawa,S., Ito,F. and Shimizu,N.
            Positional cloning of the APECD gene
            Nat. Genet. 17 (4), 393-398 (1997)
JOURNAL    98061086
MEDLINE    9398839
PUBMED
REFERENCE  2 (bases 1 to 1463)
AUTHORS    Shimizu,N.
TITLE      Direct Submision
JOURNAL    Submitted (16-AUG-1997) Nobuyoshi Shimizu, Keio University School
            of Medicine, Department of Molecular Biology, 35 Shinanomachi,
            Shinjuku-ku, Tokyo 160, Japan (E-mail:shimizu@mb.med.keio.ac.jp,
            Tel:03-3351-2370, Fax:03-3351-2370)
FEATURES   location/Qualifiers
            1..1463
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /isolate="Caucasian"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q22.3"
            /sex="male"
            /tissue_type="thymus"
            /dev_stage="3-yr-old"
            1..1463
            /gene="APECD"
            /gene="APECD"
            237..1001
            /gene="APECD"
            /note="autoimmune regulator-3"
            /codon_start=1
            /product="AIR-3"
            /protein_id="BA023993.1"
            /db_xref="GI:2696623"
            /translation="MMLVSSGAPGTQDPARNRVFPPIGMAPGVCMRPDGMGTGGG
            RISGSMAGORLSSSGTGRCSGSCFGEVALRVVHPSPVCMGVSCLCKNEDC
            AVCRGELICDQGPRAFLHCLSPPLRLIPSGRSCSLQATVQVORAEPRP
            QEPVPEPLPGIARSAGEPRCCQMTPEPCPYCMVLRVSRITLIVAVGAEVRT
            CCGLITAPLPESTGANTSPARPGRACADPAQRT"
            1449..1454
            /gene="APECD"

```

```

ORIGIN
polyA_signal
Query Match 34.0%; Score 692; DB 9; Length 1463;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1345 AGCCCCGCTGCGAGAGGCTGAGACTCCTCGGAGCTGACACCCCTACTGTGTGTGAGTCTGTA 1404
Db      772 AGCCCCGCTGCGAGAGGCTGAGACTCCTCGGAGCTGACACCCCTACTGTGTGTGAGTCTGTA 831
Qy      1405 GGGTCAGAGAACTGGCTTCTGTGTGCGGCTTGCAGGAGTGTGCGAGATGTAAGAGCT 1464
Db      832 GGGTCAGAGAACTGGCTTCTGTGTGCGGCTTGCAGGAGTGTGCGAGATGTAAGAGCT 891
Qy      1465 GCTGCGGTGTACTCACTGCGCGCGCTTCACTGAGGAGCTGCCACTTCCAGCGGAGAG 1524
Db      892 GCTGCGGTGTACTCACTGCGCGCGCTTCACTGAGGAGCTGCCACTTCCAGCGGAGAG 951
Qy      1525 CTCGCGGCTCGGAGAGAGGAGCTGCGCTGAGATCTGCTCAGAGAGAGTGAAGAGAGAGAGAG 1584
Db      952 CTCGCGGCTCGGAGAGAGGAGCTGCGCTGAGATCTGCTCAGAGAGAGTGAAGAGAGAGAGAG 1011
Qy      1585 TGTGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1644
Db      1012 TGTGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1071

```

```

Qy 1645 CACTGCACTACAGAGCCGCTCTGCACAGGATGACTGAGTCCCTTCTGAGCGAGCA 1704
Db 1072 CACTGCACTACAGAGCCGCTCTGCACAGGATGACTGAGTCCCTTCTGAGCGAGCA 1131
Qy 1705 CACCTTCATGGCATCTCTGAGTGGGCGCATTCAGAGCATGGCCGCTCCGGCGCCCTT 1764
Db 1132 CACCTTCATGGCATCTCTGAGTGGGCGCATTCAGAGCATGGCCGCTCCGGCGCCCTT 1191
Qy 1765 CCCCTCTGACCCAGATGAGCGGAGCATGACAGCTGTGATGAGAGAGTGTGAGAGAGAC 1824
Db 1192 CCCCTCTGACCCAGATGAGCGGAGCATGACAGCTGTGATGAGAGAGTGTGAGAGAGAC 1251
Qy 1825 ACCTCTCTCTGAGTCTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCCACCT 1884
Db 1252 ACCTCTCTCTGAGTCTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCCACCT 1311
Qy 1885 CTTCATGAGTCTGTGAGTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCCACCT 1944
Db 1312 CTTCATGAGTCTGTGAGTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCCACCT 1371
Qy 1945 TGGAAATTAACCCCTGCGCCCACTTCTCTACTGTGAGAGTCCCGGAGGAGCTCTCTGAGC 2004
Db 1372 TGGAAATTAACCCCTGCGCCCACTTCTCTACTGTGAGAGTCCCGGAGGAGCTCTCTGAGC 1431
Qy 2005 TGGTGAACCTACTATAAAATATAAAATTAAGCTG 2036
Db 1432 TGGTGAACCTACTATAAAATATAAAATTAAGCTG 1463

RESULT 8
LOCUS HSLTPFK 2586 bp DNA linear PRI 29-APR-1996
DEFINITION H.sapiens gene for liver type phosphofructokinase.
ACCESSION X80853
VERSION X80853.1 GI:1292864
KEYWORDS phosphofructokinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS 1
Levanon, D., Brandeis, M., Bernstein, Y. and Groner, Y.
TITLE Common promoter features in human and mouse liver type
phosphofructokinase gene
JOURNAL Biochem. Mol. Biol. Int. 35 (5), 929-936 (1995)
MEDLINE 96018039
PUBMED 7549935
REFERENCE
AUTHORS 2 (bases 1 to 2586)
Groner, Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1994) Y. Groner, The Weizmann Institute, Rehovot
76100, ISRAEL
FEATURES
source
location/Qualifiers
1..2586
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone_lib="chromosome 21 enriched genomic library
(Cytometry 7;411-417;1986)"
1..2586
/product="phosphofructokinase"
/notes="liver type"
447..616
/notes="1/2Alu"
repeat_region
/notes="1/2Alu"
repeat_region
/notes="1/2Alu"
misc_binding
/notes="Alu"
2471..2526
/bound_moiety="Sp1"
ORIGIN

```

```

Query Match 16.5%; Score 336; DB 9; Length 2586;
Best Local Similarity 100.0%; Pred. No. 5,1e-169;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1701 AGCACAACCTTGATGGCATCTCTGAGTGGGCGCATTCAGAGCATGGCCGCTCCGGCGCC 1760
Db 135 AGCACAACCTTGATGGCATCTCTGAGTGGGCGCATTCAGAGCATGGCCGCTCCGGCGCC 194
Qy 1761 CTTCCCTCTGACCCAGATGAGCGGAGCATGACAGCTGTGATGAGAGAGTGTGAGAGAGAC 1820
Db 135 CTTCCCTCTGACCCAGATGAGCGGAGCATGACAGCTGTGATGAGAGAGTGTGAGAGAGAC 254
Qy 1821 GGACAACCTCTCTGAGTCTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCC 1880
Db 255 GGACAACCTCTCTGAGTCTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCC 314
Qy 1881 ACCTCTGTGAGTCTGTGAGTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCC 1940
Db 315 ACCTCTGTGAGTCTGTGAGTGTGAGAGCGCGGCTGTGAGATCAGAGAGGAGCAGCGCC 374
Qy 1941 TGGTGAATTAACCCCTGCGCCCACTTCTCTACTGTGAGAGTCCCGGAGGAGCTCTCTGAGC 2000
Db 375 TGGTGAATTAACCCCTGCGCCCACTTCTCTACTGTGAGAGTCCCGGAGGAGCTCTCTGAGC 434
Qy 2001 TGGTGAACCTACTATAAAATATAAAATTAAGCTG 2036
Db 435 TGGTGAACCTACTATAAAATATAAAATTAAGCTG 470

RESULT 9
LOCUS AB006684 20000 bp DNA linear PRI 14-APR-2000
DEFINITION Homo sapiens APECD gene for AIRE-1, AIRE-2, AIRE-3, complete cds.
ACCESSION AB006684
VERSION AB006684.1 GI:2696618
KEYWORDS AIRE-1; AIRE-3; AIRE-2; autoimmune regulator-1; APECD; alternative
splicing.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS 1 (sites)
Nagamine, K., Peterson, P., Scott, H.S., Kudoh, J., Minoshima, S.,
Heino, M., Krohn, K.J.E., Lalioi, M.D., Mullis, P.E.,
Antonarakis, S.E., Kawasaki, K., Asakawa, S., Ito, F. and Shimizu, N.
TITLE Positional cloning of the APECD gene
JOURNAL Nat. Genet. 17 (4), 393-398 (1997)
MEDLINE 98061086
PUBMED 9398839
REFERENCE
AUTHORS 2 (bases 1 to 20000)
Shimizu, N.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1997) Nobuyoshi Shimizu, Keio University School
of Medicine, Department of Molecular Biology, 35 Shinanomachi,
Shinjuku-ku, Tokyo 160-8582, Japan
(E-mail: shimizu@mb.med.keio.ac.jp, Tel:81-3-3351-2370 (ex.2720),
Fax:81-3-3351-2370)
FEATURES
source
location/Qualifiers
1..20000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="D4G11"
/cell_line="GM130B"
/cell_type="B-lymphoblastoid"
join(4580..4838,5257..5431,5678..5833,6217..6291,
7045..7158,8357..8502,8688..8768,9795..9910,11002..11101,
11693..11875,12489..12610,13100..13202,15082..15144,
16355..16688)
/notes="AIRE-1"
4580..4838
exon

```

```
5'UTR /number=1
gene 4580..4706
CDS /gene="APECED"
join(4707..4838,5257..5431,5678..5833,6217..6291,
7045..7158,8357..8502,8688..8768,9795..9910,11002..11101,
11693..11875,12489..12610,13100..13202,15082..15144,
16355..16426)
/gene="APECED"
/note="alternative splicing"
/codon_start=1
/product="AIRE-1"
/protein_id="BAA23990.1"
/db_xref="GI:2696619"
/translation="MATDAALRLRLRLHRTETAVAVDSAPFLHALLADHDVPERKQ
ETLHLKEKGCQAFHALSLWLTQDSTALIDFWVLFKDYNLERGLQPLDSFPK
DVDLSQPRKGRKPPKALVPPRLPTKRKASEEAPALPTREGIASPSQLKA
KPKPKESSAEQQLPLGNGIQTMASVQRAVAMSSGDVPAAGAVEGILIQVFESG
GSKCCTIOVGEEYTPSKFEDSGSKMKARSSGPRPLVAKAGACGAPGGBARLGGQ
GSVPALALPSDPOLHOKNEDRCVCRGSELICCDGCPRAFLACLSPRLRETSCT
WRSSCLOATVOEVQPRABEPRPQEPVETPLPPGLRSAGEVRGPGEPLIGMDTTL
VYKHLPPPSAAPLPGLDSSALHPLLCVGEQQLNLAPEGARCGVCGDGTDLVLRTHCA
AAPHRCHFPAGTSRPGTGLRCRSCSGDVTAPAVEGVLAAPSPARLAPGPAKDDTASHE
PALHRDLESLSEHTFDGILQWAIQSMARPAAPPS"
exon 5257..5431
/gene="APECED"
/number=2
exon 5678..5833
/gene="APECED"
/number=3
exon 6217..6291
/gene="APECED"
/number=4
exon 7045..7158
/gene="APECED"
/number=5
exon 8357..8502
/gene="APECED"
/number=6
exon 8688..8768
/gene="APECED"
/number=7
exon join(8965..9148,9455..9910,11002..11101,11693..11875,
12489..12610,13100..13202,15082..15144,16355..16688)
/note="AIRE-2"
join(8965..9148,9455..9910,11002..11101,11693..11725,
11808..11875,12489..12610,13100..13202,15082..15144,
16355..16688)
/note="AIRE-3"
join(8965..9148,9455..9506)
/gene="APECED"
exon 8965..9148
/gene="APECED"
/number=1
exon 9455..9910
/gene="APECED"
/number=8
CDS join(9507..9910,11002..11101,11693..11875,12489..12610,
13100..13202,15082..15144,16355..16426)
/gene="APECED"
/note="alternative splicing
autoimmune regulator-2"
/codon_start=1
/product="AIRE-2"
/protein_id="BAA23991.1"
/db_xref="GI:2696620"
/translation="MWLVSSGAPGTQPARNRVFPPIGMARPGVCMRPDNGTGGQG
RISGSGMAGQRLSSGTQRCSGSKKEVALRLVLPSPVCMGVSCLCKNEDRC
AVCRDGEILICDDGCPRAFLACLSPRLREISGTWRSSCLOATVOEVQPRABEPR
QEPVETPLPPGLRSAGEVRGPGEPLIGMDTTLVYKHLPPPSAAPLPGLDSSALH
PLLCVGEQQLNLAPEGARCGVCGDGTDLVLRTHCAAPHRCHFPAGTSRPGTGLR
SCSGDVTAPAVEGVLAAPSPARLAPGPAKDDTASHEPALHRDLESLSEHTFDGILQ
WAIQSMARPAAPPS"
```

```
CDS join(9507..9910,11002..11101,11693..11875,12489..12610,13100..13137)
/gene="APECED"
/note="alternative splicing"
/codon_start=1
/product="AIRE-3"
/protein_id="BAA23992.1"
/db_xref="GI:2696621"
/translation="MWLVSSGAPGTQPARNRVFPPIGMARPGVCMRPDNGTGGQG
RISGSGMAGQRLSSGTQRCSGSKKEVALRLVLPSPVCMGVSCLCKNEDRC
AVCRDGEILICDDGCPRAFLACLSPRLREISGTWRSSCLOATVOEVQPRABEPR
QEPVETPLPPGLRSAGEVRGPGEPLIGMDTTLVYKHLPPPSAAPLPGLDSSALH
PLLCVGEQQLNLAPEGARCGVCGDGTDLVLRTHCAAPHRCHFPAGTSRPGTGLR
SCSGDVTAPAVEGVLAAPSPARLAPGPAKDDTASHEPALHRDLESLSEHTFDGILQ
WAIQSMARPAAPPS"
exon 9795..9910
/gene="APECED"
/number=8
exon 11002..11101
/gene="APECED"
/number=9
exon 11693..11875
/gene="APECED"
/number=10
exon 11693..11725
/gene="APECED"
/note="10a"
/number=10
exon 11808..11875
/gene="APECED"
/note="10b"
/number=10
exon 12489..12610
/gene="APECED"
/number=11
exon 13100..13202
/gene="APECED"
/number=12
exon join(13138..13202,15082..15144,16355..16688)
15082..15144
/gene="APECED"
/number=13
exon 16355..16688
/number=14
16427..16688
ORIGIN
3'UTR
Query Match 16.5%; Score 336; DB 9; Length 20000;
Best Local Similarity 100.0%; Pred. No. 3.8e-169;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1701 AGCACACCTTGATGAGCATCTCTGCGAGTGGGCCATTCAGAGCATGGCCCTCGGCGGCC 1760
DB 16353 AGCACACCTTGATGAGCATCTCTGCGAGTGGGCCATTCAGAGCATGGCCCTCGGCGGCC 16412
QY 1761 CCTTCCTCTGAGCCCGAGATGGCCGGAGCATGAGCTCTGATGAGAGTGTGAGAA 1820
DB 16413 CCTTCCTCTGAGCCCGAGATGGCCGGAGCATGAGCTCTGATGAGAGTGTGAGAA 16472
QY 1821 GGACACCTCTTCTCTCACTCTGGAAGCCGGCGCTGGATCAAGAAAGGGACAGCGCC 1880
DB 16473 GGACACCTCTTCTCTCACTCTGGAAGCCGGCGCTGGATCAAGAAAGGGACAGCGCC 16532
QY 1881 ACCTTTTTCAGTGTGTCGGCTGTAAACAGCTCTGTGTTTTCGGGACACAGCCCTCATG 1940
DB 16533 ACCTTTTTCAGTGTGTCGGCTGTAAACAGCTCTGTGTTTTCGGGACACAGCCCTCATG 16592
QY 1941 TGCTGGAATTAACCTGCGCCCACTTCTACTCTGGAAGTCCCGGAGACCTCTCT 2000
DB 16593 TGCTGGAATTAACCTGCGCCCACTTCTACTCTGGAAGTCCCGGAGACCTCTCT 16652
QY 2001 TGCTGTGACCTACTAAATATATAAATTAGCTG 2036
DB 16653 TGCTGTGACCTACTAAATATATAAATTAGCTG 16688
```



[illegible]

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
FEATURES  
SOURCE

Homo sapiens (human)  
Homo sapiens  
Bakayocsa, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1  
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,  
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,  
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,  
Pollay,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,  
Peterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,  
Zimmermann,W., Rosenthal,A., Kudof,J., Shibuya,K., Kawasaki,K.,  
Aakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Miyayama,S.,  
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G.,  
Hornischer,K., Barand,P., Scharfe,M., Schoen,O., Desario,A.,  
Reichelt,J., Kauer,G., Blocker,H., Ramser,J., Beck,A., Klages,S.,  
Hennig,S., Riesselmann,L., Daand,E., Wehrmeyer,S., Borzym,K.,  
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and  
Yaspo,M.L.  
The DNA sequence of human chromosome 21  
Nature 405 (6784), 311-319 (2000)  
20289799  
10830953  
2 (bases 1 to 340000)  
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,  
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,  
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,  
Pollay,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,  
Peterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,  
Zimmermann,W., Rosenthal,A., Kudof,J., Shibuya,K., Kawasaki,K.,  
Aakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Miyayama,S.,  
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G.,  
Hornischer,K., Barand,P., Scharfe,M., Schoen,O., Desario,A.,  
Reichelt,J., Kauer,G., Blocker,H., Ramser,J., Beck,A., Klages,S.,  
Hennig,S., Riesselmann,L., Daand,E., Wehrmeyer,S., Borzym,K.,  
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and  
Yaspo,M.L.  
Direct Submission  
Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing  
Consortium: \* RIKEN Genomic Sciences Center, Human Genome Research  
Group \* Institute of Molecular Biotechnology, Genome Analysis  
Keio University School of Medicine, Dept. of Molecular Biology \*  
GfP, Dept. of Genome Analysis \* Max-Planck Institute for Molecular  
Genetics (addresses see below)  
On May 30, 2000 this sequence version replaced gi:7717429.  
The chromosome 21 mapping and sequencing consortium consisting of  
\* RIKEN Genomic Sciences Center, Human Genome Research Group, \*  
Sagamihara 228-8555, Japan,  
\* e.mail: hattori@gsc.riken.go.jp  
\* URL: http://hgp.gsc.riken.go.jp/  
and  
\* Institute of Molecular Biotechnology, Genome Analysis, \*  
Beutenbergstrasse 11, D-07745 Jena, Germany,  
\* e.mail: gscg-submit@genome.imb-jena.de  
\* URL: http://genome.imb-jena.de/  
and  
\* Keio University School of Medicine, Molecular Biology, \* Tokyo  
160-8582, Japan,  
\* e.mail: nshimizu@db.med.keio.ac.jp  
\* URL: http://www.med.keio.ac.jp/  
and  
\* GfP, Dept. of Genome Analysis,  
Mascheroder Weg 1, D-38124 Braunschweig, Germany, \* e.mail:  
info.genome@gf.de  
\* URL: http://genome.gbf.de/  
and  
\* Max-Planck Institute for Molecular Genetics,  
Innestrasse 73, D-14195 Berlin, Germany,  
\* e.mail: info-chr21molgen.mpg.de  
\* URL: http://chr21.rz-berlin.mpg.de/  
AL163299. Submitted (10-Apr-2000).  
Location/Qualifiers  
1. 340000  
/organism="Homo sapiens"

```

source
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    .1. .17637
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="KUD4G11, 5' partial"
    /clone_1fb="KU21-D Cosmid library",
    /note="Accession No. AP001060"
    16296. .32142
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="KUD6A4"
    /clone_1fb="KU21-D Cosmid library"
    /note="Accession No. AP001061"
    25925. .62855
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="Q05B10"
    /clone_1fb="L121NCO2-Q Cosmid library"
    /note="Accession No. AP001062"
    56250. .101561
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="KUD99F9"
    /clone_1fb="KU21-D Cosmid library"
    /note="Accession No. AP001063"
    97304. .135091
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="KUD107C4"
    /clone_1fb="KU21-D Cosmid library"
    /note="Accession No. AP001064"
    109897. .225790
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="XB68A7"
    /clone_1fb="Keio BAC library"
    /note="Accession No. AP001065"
    212299. .251501
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="KUD11H9"
    /clone_1fb="KU21-D Cosmid library"
    /note="Accession No. AP001066"
    230165. .2340000
source
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q22.3"
    /clone="KB1399C7, 3' partial"
    /map="21q22.3"

```

	source	/clone_lib="Keio BAC library" /note="Accession No. AP001067" 280806. .>340000 organism="Homo. sapiens" mol_type="genomic DNA" db_xref="taxon:9606" chromosome="21" map="21q22.3"
gene		/clone="P314N7, 3' partial" /clone_1ib="RPC11.3-5 PAC library" /note="Accession No. AJ011930" 1508. .13840 /gene="AIRE"
mRNA		join(1508. .1759,2178. .2352,2599. .2754,3138. .3212, 3366. .4079,5278. .5423,5609. .5689,6716. .6831,7923. 8614. .8796,9410. .9531,10022. .10124,12004. .12066, 13277. .13840) /gene="AIRE"
exon		/product="autoimmune regulator (APECED protein)" /note="Accession No. Z97990" 1508. .1759 /gene="AIRE"
5'UTR		/number=1 1508. .1627 /gene="AIRE"
CDS		join(1628. .1759,2178. .2352,2599. .2754,3138. .3212, 3366. .4079,5278. .5423,5609. .5689,6716. .6831,7923. 8614. .8796,9410. .9531,10022. .10124,12004. .12066, 13277. .13348) /gene="AIRE"
exon		/note="Accession No. Z97990" /codon_start=1 /product="autoimmune regulator (APECED protein)" /protein_id="BAA95560.1" /db_xref="GI:7768776"
exon		/translation="MATDAALRLRLRHRTETAVAVDAPFLHALADHDVDEDKQ ETLHKKEGCGQAOFHALISWLTQDSTAILDFWRVLFDYNLERYGRLOPIIDSEPK DVDSLCKPRKRPAPVPKALVPPPLPTKRKAEEARPAAPALTRPGTASPSOLKA KPPESSAESBOQLPLNGIOTMSASVORAVAMSGDVGAAGAVEGLIIQQVFESG GSVKPIIOGEFYPSPKFEDSGSKNKASSSGPKPLVRAKOGAAPGGEGEARLGQ GSVPKPLALPSDPOLHOKNEDECACNRDGELICCDGCPRAHLACTLSPRIIPST WRSCSLATIVDEVOPRAEERFOEPVEITPLPGRSAGEEVRRGPGEPLAMDITTL VYKHLPAPSAAPLPGLDSSALHPLLCVBPBEQONILAPARCVCVDGDIVLCITCA AAFWHRCHFPAGTSRRPGTLRCRSCSDVTTPAPEGVLLPSPARLAPFAKDPTASH PALHDDLESLSBHFTFGILLQMAIQSMARPAPAFPS"
exon		2178. .2352 /gene="AIRE" /number=2 2599. .2754
Query Match	16.5%;	Score 336; DB 9; Length 340000;
Best Local Similarity	100.0%;	Pred. No. 2.5e-169;
Matches 336;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Db	1701	AGCACACCTTCAGATGCATCTCGACGTGGCCATCCAGAAGCATGCCCGTCCGGCGGCC 1760
Qy	13275	AGCACAACCTTGATGGCATCTCGACGTGGCCATCCAGAAGCATGCCCGTCCGGCGGCC 13334
Db	1761	CTTTCCCTCTCGACCCCCAGATGCGCGGAGCATGCAGCTCTGTATGAGAGTGTCTGAGAA 1820
Qy	13335	CTTTCCCTCTCGACCCCCAGATGCGCGGAGCATGCAGCTCTGTATGAGAGTGTCTGAGAA 13394
Db	1821	GAGACACTCTCTCGACGCTCGGAGCGGCGCGGTGGAGATCAAGAGGGAGCAGCGCC 1880
Qy	13395	GAGACACTCTCTCGACGCTCGGAGCGGCGCGGTGGAGATCAAGAGGGAGCAGCGCC 13454
Db	1881	ACCTCTGTCACTGCTCGGCTGTAAAACAGCTCTGTATTCTGGGGAACACAGCATCATG 1940
Qy	13455	ACCTCTGTCACTGCTCGGCTGTAAAACAGCTCTGTATTCTGGGGAACACAGCATCATG 13514
Db	1941	TGCTCGGAATTAACCTGCGCCCACTTCTTAATCTTGAAAGTCCCCGGAGGCTTCTCT 2000
Qy	13515	TGCTCGGAATTAACCTGCGCCCACTTCTTAATCTTGAAAGTCCCCGGAGGCTTCTCTCT 13574

Qy 2001 TCCTGTGACCTACTAAAAATATAAATTAGCTG 2036  
|||||  
Db 13575 TCCTGTGACCTACTAAAAATATAAATTAGCTG 13610

RESULT 12  
HSAJ9610  
LOCUS HSAJ9610 36284 bp DNA linear PRI 01-SEP-2000  
DEFINITION Homo sapiens AIRE gene.  
ACCESSION AJ009610  
VERSION AJ009610.1 GI:3392939  
KEYWORDS Aire gene.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
1 Aaltonen,J., Bjreses,P., Perheentupa,J., Horelli-Kuitunen,N.,  
Palot,A., Pelttonen,L., Lee,Y.S., Francis,F., Hennig,S., Thiel,C.,  
Lehrach,H. and Yaspo,M.L.  
An autoimmune disease, APECED, caused by mutations in a novel gene  
featuring two PHD-type zinc finger domains  
Nat. Genet. 17, 399-403 (1997)  
2 (bases 1 to 36284)  
Lee,Y.S., Francis,F., Hennig,S., Thiel,C., Reinhard,R., Lehrach,H.  
and Yaspo,M.L.  
Direct Substitution  
Submitted (17-JUL-1998) Steffen Hennig, MPIHG, Abt. Lehrach, Max  
Planck Institut fuer Molekulare Genetik, Innestrasse 73, Berlin,  
14195, Germany  
Related cDNA sequence: Z97990.  
FEATURES  
location/Qualifiers  
1..36284  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="q22.3"  
/clone\_11b="Lawrence Livermore Cosmid 22G11"  
199..323  
repeat\_region  
/note="L1MC3"  
complement(779..1090)  
/note="AluY"  
repeat\_region  
complement(1101..1198)  
/note="AluSc"  
complement(1245..1533)  
/note="AluSg"  
complement(1537..1678)  
/note="AluY"  
complement(1688..1823)  
/note="AluSc"  
complement(1824..2106)  
/note="AluSg"  
complement(2112..2196)  
/note="AluSg/X"  
complement(2208..2234)  
/note="(GAAA)n"  
complement(2235..2532)  
/note="AluSx"  
complement(3109..3382)  
/note="AluY"  
complement(5332..5634)  
/note="AluSg"  
complement(5722..5820)  
/note="(TAAA)n"  
6217..6515  
repeat\_region  
/note="AluSg"  
6531..6761  
/note="AluSc"  
complement(6805..7106)  
/note="AluSp"

repeat\_region complement(7107..7409)  
/note="AluSp"  
7445..7729  
repeat\_region  
/note="AluSg"  
complement(7757..7798)  
/note="(GAAA)n"  
complement(7799..8101)  
/note="AluSx"  
10206..10439  
repeat\_region  
/note="L1"  
10443..10742  
/note="AluSx"  
complement(10918..11215)  
/note="AluSx"  
complement(11501..11878)  
/note="MSTB"  
complement(11993..12079)  
/note="MSTB"  
12193..12880  
repeat\_region  
/note="L1"  
complement(12887..13178)  
/note="AluSg"  
13289..13582  
/note="AluSx"  
13585..13738  
/note="L1"  
complement(13751..14060)  
/note="L1MB4"  
complement(14061..14357)  
/note="AluY"  
complement(14360..14720)  
/note="L1MB4"  
complement(14722..15018)  
/note="AluY"  
complement(15019..15394)  
/note="L1MB4"  
complement(15398..15700)  
/note="AluSg"  
complement(15704..15864)  
/note="L1MB4"  
15866..16182  
/note="L1"  
16726..16931  
/note="AluSg/X"  
complement(16933..17084)  
/note="MLT1A1"  
complement(17090..17375)  
/note="AluSx"  
complement(17379..17530)  
/note="MLT1A1"  
17663..17959  
/note="AluY"  
18738..18920  
/note="MIR"  
complement(19498..20418)  
/note="LTR5"  
complement(21877..21968)  
/note="MIR"  
complement(22524..22642)  
/note="GC-rich"  
22529..34854  
/gene="AIRE"  
22529..22779  
/gene="AIRE"  
/number=1  
join(22648..22779,23198..23372,23619..23774,24158..24232,  
24986..25099,26298..26443,26629..26709,27336..27851,  
28943..29042,29633..29815,30428..30459,31040..31142,  
33022..33084,34291..34362)  
/gene="AIRE"  
/codon\_start=1  
/protein\_id="CAA08759.1"  
/db\_xref="GI:3392940"

```
/db_xref="GDB:043918"
/db_xref="Swiss-Prot:043918"
/translation="MATDPAALRLRLHTEIAVAVDSAPFLHLALADHDVPEDEKQ
ETLHLKEKGCPOAFHALSLTLDSTALIDFWVLFDVNLERGRLOPLDSEPK
KVDLSQPRKGRPAVPKALVPPRLPTRKASBEMAPALPRGTASPGSOLKA
PPDKPSSASBOQRLPLGNGIQTMASVQRAVMSGDVPGARAVEGLIIQVSESG
GSKKICVNGEFTYSKPEDSGSKNKASSSGPKPLVAKAGOGAPGGARIGQO
GSVPAIALPSDIPOLHKNEDECAVCRDGEILICDQCPRAFLACLSPPLEISGT
WACSSCLOATVDEVQPRAEPRPQEPVETPLPPIGRSAGEVRCPPGPIAGMTTL
VKHLPAIPSAAPLPGDSSALHPLLCVPEEQNLAPGARGVWGTLCRSCSDVT
PAVGVLAIPSAAPLPAKPDASHPALHRLDLESLSHHTPGILQMAIQSMAR
PAAPPS"
23198..23372
/gene="AIRE"
/exon
number=2
23619..23774
/gene="AIRE"
number=3
24158..24232
/gene="AIRE"
number=4
24986..25099
/gene="AIRE"
number=5
complement(25795..25875)
repeat_region
note="MULTI1"
complement(25956..26114)
repeat_region
note="MULTI1"
26298..26443
/gene="AIRE"
number=6
26629..26709
/gene="AIRE"
number=7
27736..27851
/gene="AIRE"
number=8
28943..29042
/gene="AIRE"
number=9
29633..29815
/gene="AIRE"
number=10
30428..30549
/gene="AIRE"
number=11
complement(30872..30983)
repeat_region
note="POLY G"
31040..31142
/gene="AIRE"
number=12
33022..33084
/gene="AIRE"
number=13
34291..34854
/gene="AIRE"
number=14
34600..34787
note="AluSg"
complement(34793..35417)
repeat_region
note="L1MB3"
complement(35423..35747)
repeat_region
Query Match 15.8%; Score 322; DB 9; Length 36284;
Best Local Similarity 100.0%; Pred. No. 1.2e-161;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1715 GGCATCTCTCAGTGGGCATTCAGAGCATGGCCCGTCCGGCGCCCTCTCCCTCTCTCA 1774
DB 34303 GGCATCTCTCAGTGGGCATTCAGAGCATGGCCCGTCCGGCGCCCTCTCCCTCTCTCA 34362
QY 1775 CCCGAGATGGCGGAGCATGCACTGTGAGAGAGTCTGAGAGAGACACTCTCTCC 1834
```

```
DB 34363 CCCGAGATGGCGGAGCATGCACTGTGAGAGAGTCTGAGAGAGACACTCTCTCC 34422
QY 1835 TAGATCTCTGGAAGCCGCGCGCTGGGATCAAGAGGGGACAGCCGCACTCTTGTCACTG 1894
DB 34423 TAGATCTCTGGAAGCCGCGCGCTGGGATCAAGAGGGGACAGCCGCACTCTTGTCACTG 34482
QY 1895 CTCGGCTGTAACAGCTCTGTGTTTCTGGGGGACACAGGCATCATGTGCTGGAATTAA 1954
DB 34483 CTCGGCTGTAACAGCTCTGTGTTTCTGGGGGACACAGGCATCATGTGCTGGAATTAA 34542
QY 1955 ACCCTGCCCACTTCTTACTCTGAGAGTCCCGGAGAGCTTCTTGGCTGTGACTTA 2014
DB 34543 ACCCTGCCCACTTCTTACTCTGAGAGTCCCGGAGAGCTTCTTGGCTGTGACTTA 34602
QY 2015 CTAATAATTAATAATTAGCTG 2036
DB 34603 CTAATAATTAATAATTAGCTG 34624
```

```
RESULT 13
AC003656_5
WPCOMMENT
Sequence split into 8 fragments LOCUS AC003656, Accession AC003656
Fragment Name Begin End
AC003656_1 1 110000
AC003656_2 100001 310000
AC003656_3 200001 310000
AC003656_4 300001 410000
AC003656_5 400001 510000
AC003656_6 500001 610000
AC003656_7 600001 710000
AC003656_8 700001 753000
Continuation (6 of 8) of AC003656 from base 500001 (AC003656 Homo sapiens clone PI C124G)
Query Match 13.2%; Score 269; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 3.9e-133;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGACCGGGAGACGGGGCGGCGCCACACGCGCGCGGAGGCCACAGCCCGCGGAGCC 60
DB 47488 AGACCGGGAGACGGGGCGGCGCCACACGCGCGCGGAGGCCACAGCCCGCGGAGCC 47547
QY 61 CGAGGCGAAGCGAGGGGCGTCCAGTCCCGGGACCAACCGCGCGCGCGCGCGCGCGG 120
DB 47548 CGAGGCGAAGCGAGGGGCGTCCAGTCCCGGGACCAACCGCGCGCGCGCGCGCGCGG 47607
QY 121 TCCCGCGCCACCCCATGCGACGAGCGCGCGCTAGCGCGCTTCTGAGGCTGACCG 180
DB 47608 TCCCGCGCCACCCCATGCGACGAGCGCGCGCGCTAGCGCGCTTCTGAGGCTGACCG 47667
QY 181 CAGGAGATCGGCTGCGCGCTGAGACAGGCGCTTCCCATGCTGACAGCGGCTGACCA 240
DB 47668 CAGGAGATCGGCTGCGCGCTGAGACAGGCGCTTCCCATGCTGACAGCGGCTGACCA 47727
QY 241 CGAGGTGTCCTCCGAGCAAGTTTCAGG 269
DB 47728 CGAGGTGTCCTCCGAGCAAGTTTCAGG 47756
```

```
RESULT 14
AL954247
LOCUS AL954247 143192 bp DNA linear PRI 16-OCT-2003
DEFINITION Pan troglodytes chromosome 22 clone CH251-479113 map 22q22.3,
complete sequence.
ACCESSION AL954247
VERSION AL954247.2 GI:37606037
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1 (bases 1 to 143192)
```

CONSTRM The Chimpanzee Chromosome 22 Sequencing Consortium  
 TITLE DNA sequence of chimpanzee chromosome 22 and its evolutionary  
 JOURNAL implications  
 REFERENCE unpublished  
 AUTHORS 2 (bases 1 to 143192)  
 Kube, M., Sudbrak, R., Mueller, I., Thiel, J., Klages, S., Borzym, K.,  
 Heilmann, K., Gimmel, V., Beck, A., Ben Kahla, A., Lehrach, H.,  
 Yaspo, M.L. and Reinhardt, R.  
 Submitted (18-DEC-2002)  
 Direct Submission  
 TITLE The Chimpanzee Chromosome 22 Sequencing Consortium consists of:  
 JOURNAL \*Chinese National Human Genome Center at Shanghai,  
 COMMENT Shanghai, China;  
 \*GDF, Dept. of Genome Analysis, Braunschweig, Germany;  
 \*Institute of Molecular Biotechnology, Jena, Germany;  
 \*KIBB Genome Research Center, Daejeon, Korea;  
 \*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;  
 \*National Institute of Genetics, Mishima, Japan;  
 \*National Yang Ming University Genome Research Center, Taipei,  
 Taiwan;  
 \*RIKEN Genomic Sciences Center, Yokohama, Japan.  
 ----- Genome Center  
 Center: Max-Planck-Institute for Molecular Genetics  
 Center code: MPIMG  
 ----- Project Information  
 Center clone name: CH251-479113  
 ----- Summary Statistics  
 Sequencing vector: pUC18; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 143190 bases at least Q40  
 Consensus quality: 143192 bases at least Q30  
 Consensus quality: 143192 bases at least Q20  
 Quality coverage: 25.4  
 -----  
 This sequence was finished as follows unless otherwise noted: all  
 regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. -----  
 Neighboring sequence information:  
 This clone is overlapped by  
 PTB-051003  
 PTB-061A04  
 PTB-061A04  
 -----  
 SOURCE INFORMATION:  
 The CHORI-251 chimpanzee BAC library was prepared from DNA isolated  
 from the blood of a single male chimpanzee using published  
 protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the  
 chimpanzee ('Clint') was obtained from the Yerkes Primate Center in  
 Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,  
 Kazuhiro Osoegawa, Evan Richter & Pieter J de Jong. The library  
 characteristics are described at  
 http://www.chori.org/bacpac/chimpanzee251.htm. The clone may be  
 obtained from Pieter J. de Jong and coworkers  
 (http://www.chori.org/bacpac). VECTOR: PTARAC2.1  
 IMPORTANT: This sequence is not the entire insert of clone  
 CH251-479113. It may be shorter because we sequence  
 overlapping  
 sections only once, except for a short overlap.  
 -----  
 Location/Qualifiers  
 1..143192  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /chromosome="22"  
 /map="22q22.3"  
 /clone="CH251-479113"  
 <1..14090  
 /organism="Pan troglodytes"

source  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-051003"  
 /note="overlapping clone"  
 141389..>143192  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-061A04"  
 /note="overlapping clone"  
 ORIGIN  
 Query Match 9.6%; Score 195; DB 9; Length 143192;  
 Best Local Similarity 100.0%; Pred. No. 3e-93;  
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 source  
 1840 CCTGGAAGCGCGCCGGCTGGATCAAGAGGGAGACGCGCCTCTTGTAGTGTGG 1899  
 DB 45529 CCTGGAAGCGCGCCGGCTGGATCAAGAGGGAGACGCGCCTCTTGTAGTGTGG 45588  
 QY 1900 CTGTAACAGCTCTGTTCTGGGAGACACGACCATCATGTGCTGAAATTAACCT 1959  
 DB 45589 CTGTAACAGCTCTGTTCTGGGAGACACGACCATCATGTGCTGAAATTAACCT 45648  
 QY 1960 GCCCACTTCTACTCTGGAAGTCCCGGAGAGCTCTTCTGCTGACTTAATA 2019  
 DB 45649 GCCCACTTCTACTCTGGAAGTCCCGGAGAGCTCTTCTGCTGACTTAATA 45708  
 QY 2020 AATATATAATTAATGC 2034  
 DB 45709 AATATATAATTAATGC 45723  
 RESULT 15  
 LOCUS CQ541728 60 bp DNA linear PAT 30-JAN-2004  
 DEFINITION Sequence 11363 from Patent WO0210449.  
 ACCESSION CQ541728  
 VERSION CQ541728.1 GI:41507992  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 Shoshan, A., Wasserman, A., Mintz, E., Mintz, L. and Paigler, S.  
 TITLE Oligonucleotide library for detecting rna transcripts and splice  
 variants that populate a transcritpome  
 JOURNAL Patent: WO 0210449-A 11363 07-FEB-2002;  
 Compugen Inc. (US)  
 FEATURES  
 source 1..60  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 2.9%; Score 60; DB 6; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-20;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1894 GCTCGGCTGTAACAGCTCTGTGTTCTGGGAGACACGACCATCATGTGCTGAAATTA 1953  
 DB 1 GCTCGGCTGTAACAGCTCTGTGTTCTGGGAGACACGACCATCATGTGCTGAAATTA 60  
 RESULT 16  
 LOCUS BV184386 101 bp DNA linear STS 10-JUN-2004  
 DEFINITION egm144551 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
 tagged site.  
 ACCESSION BV184386  
 VERSION BV184386.1 GI:48022640



JOURNAL	Submitted (17-FEB-1999)	Pathology, University of Florida, 1600 SW Archer Road
FEATURES	Location/Qualifiers	
source	1..1656	/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="SJL"
		/db_xref="taxon:10090"
		/chromosome="10"
		/map="between D10Mit31 and D10Mit10"
gene	1..>1656	/gene="Alre"
CDS	1..>1656	/gene="Alre"
		/function="transcription factor"
		/codon_start=1
		/product="autoimmune regulator"
		/protein_id="AFJ6481.1"
		/db_xref="GI:7108573"
		/translation="MAGDGLMLRRLRLHRTETIAVAIDSAFPLHLADHDVPEDEF OETLRLEKEGCPQAFALLSVLTBDSGAILDFPWRILFPDYMLERSYSLIIDGP KVDIVNKRGRDLAIGPKAAVLPPPTGRKLAEPRATLPATLAKSSVSPSHLK TKRPKPDGRLSEOHLPICNGIOTMAISVORATAVGGAPRGRAVEGILLIOVPES GRSKCTIQVGGEYTTPNKFPEDSGNLKNKARSSSLKPYVRAGAQTITGRBRQKG QQCGVCPPLPSIPSEPYNQNDNECAVCHDGSELICDDGCPRAFHLACLSPIQELTS GLWRCSPPLSRIFSEBPALPLEPSALCPLISAGNEGPGPAAPSARCSCVCGGTVELCAHCQA AVTYVYLHPAPAIPLEPSALCPLISAGNEGPGPAAPSARCSCVCGGTVELCAHCQA AATHMCCLPPTAAPRGTNIRCKSCASGSTPTGTEAAPTGSRAVPTGSPRAPFLAKVGSDS ASHDPVLYHRDLESLLNHSFDDILGMALGMSRLPAETPFSS"
ORIGIN		
Query Match	1.6%; Score 32;	DB 10; Length 1656;
Best local Similarity	100.0%;	Pred. No. 4.7e-05;
Matches 32; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Oy	1484	GCGCGTCCCTTCACACTGGCGCTGCCACTTCCC 1515
Db	1348	GCGCGTGCTTCCACTGCGCTGCCACTTCCC 1379
RESULT 21		
LOCUS	AF128773	1656 bp mRNA linear ROD 26-OCT-2001
DEFINITION	Mus musculus strain NOD autoimmune regulator (Aire) mRNA, partial cds.	
ACCESSION	AF128773	
VERSION	AF128773.1	GI:7108574
KEYWORDS	Mus musculus (house mouse)	
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE AUTHORS	1 (bases 1 to 1656) Shi,J.D., Wang,C.Y., Marron,M.P., Ruan,Q.G., Huang,Y.Q., Detter,J.C. and She,J.X. Chromosomal localization and complete genomic sequence of the murine autoimmune regulator gene (Aire)	
TITLE	Autoimmunity 31 (1), 47-53 (1999)	
JOURNAL	Medline 20059142	
PUBMED	10593569	
REFERENCE	2 (bases 1 to 1656) Shi,J.D., Wang,C.Y., Marron,M.P., Ruan,Q.G., Huang,Y.Q., Detter,J.C., Davoodi-Semirioni,A. and She,J.X. Complete genomic sequence, gene structure and localization of the mouse Aire gene	
AUTHORS	Unpublished	
TITLE	3 (bases 1 to 1656) Ruan,Q.G., Wang C.Y., Shi,J.D. and She,J.X. Direct Submission	
JOURNAL	Submitted (17-FEB-1999)	Pathology, University of Florida, 1600 SW Archer Road
FEATURES	Location/Qualifiers	
source	1..1656	



/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NOD"  
 /db\_xref="taxon:10090"  
 /chromosome="10"  
 /map="between D10Mit31 and D10Mit10"  
 1. .1656  
 /gene="Aire"  
 1. .1656  
 /gene="Aire"  
 /function="transcription factor"  
 /codon\_start=1  
 /product="autoimmune regulator"  
 /protein\_id="AAP36482.1"  
 /db\_xref="GI:7108575"  
 /translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHNLADHDVPEDKF  
 QETLRLKEEGCPQAFHALLSWLTRDSGLDFWRLFKDYNLERYSLHSILGFP  
 KDVDLNSRKGRKPLAGPKAAVLPRPPTRKALREPRATPLASKVSFSGSLK  
 TKPKPKDGNLESQHLPLGNGIOTMAASVQRAVTVASGDPVGTGAVGILLIOVRES  
 GSKKCIQVGGFPTPNKEDPSGNLKKKASGSLKPYRAKAGQVITPGRDEQVQ  
 OCGVPPPLPSPEPVONKNEDECAVCHDGSELICDGCPRAFHLACTLSPLOEIPS  
 GIMRSCCTIQGRVQNLISQPEVSRPELPAETPIIVGLRSASEKTRGSRBELKASDA  
 AVTVNLAPHPAPALPESALCPILSAGNEGRPGAPARSVCGDGTEVLRCACHCA  
 AAFHWRCHPPTAARPGTNLRCKSCSADSTPTGPGEAIVPTSGPRPAPGLAKVGDD  
 ASHDPVLRHLDLESLNLNHSFDGILLQMAIQMSRPLAETPPFS"

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1656;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACCTGCGGCTGCCACTTCCC 1515  
 ||||||||||||||||||||||||||||||||||||  
 DB 1348 GCCGCTGCTTCACCTGCGGCTGCCACTTCCC 1379

RESULT 22  
 A97293 1659 bp DNA linear PAT 26-JAN-2000  
 LOCUS A97293 Sequence 10 from Patent WO9918197.  
 DEFINITION A97293  
 VERSION A97293.1 GI:6780676  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 REFERENCE 1 (bases 1 to 1659)  
 YASPO,M. and Lehrach,H.  
 NUCLEIC ACID MOLECULE ENCODING A (POLY)PEPTIDE CO-SEGREGATING IN  
 MUTATED FORM WITH AUTOIMMUNE POLYENDOCRINOPATHY CANDIDIASIS  
 ECTODERMAL DYSTROPHY (APECED)  
 PATENT: WO 9918197-A 10 15-APR-1999;  
 JOURNAL MAX PLANCK GESSELLSCHAFT (DE); YASPO MARIE LAURE (DE)  
 FEATURES  
 source 1. .1659  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

## ORIGIN

Query Match 1.6%; Score 32; DB 6; Length 1659;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACCTGCGGCTGCCACTTCCC 1515  
 ||||||||||||||||||||||||||||||||||||  
 DB 1348 GCCGCTGCTTCACCTGCGGCTGCCACTTCCC 1379

RESULT 23  
 MMU132243 1659 bp mRNA linear ROD 24-APR-1999  
 LOCUS MMU132243  
 DEFINITION Mus musculus mRNA for Aire protein.

ACCESSION AJ132243  
 VERSION AJ132243.1 GI:4456674  
 KEYWORDS Aire gene; Aire protein.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1  
 AUTHORS Blechschmidt,K., Schweiger,M., Wetz,K., Poulsen,R.,  
 Christensen,H.M., Rosenthal,A., Lehrach,H. and Yaspo,M.L.  
 TITLE The mouse Aire gene: comparative genomic sequencing, gene  
 organization, and expression  
 JOURNAL Genome Res. 9 (2), 158-166 (1999)  
 MEDLINE 99148139  
 PUBMED 10022980  
 REFERENCE 2 (bases 1 to 1659)  
 AUTHORS Karin,B.  
 TITLE Direct Submision  
 JOURNAL Submitted (14-JAN-1999) Karin B., Max Planck Institute for  
 Molecular Genetics, Ihnestrasse 73, Berlin D-14195, Germany  
 FEATURES  
 source 1. .1659  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129"  
 /db\_xref="taxon:10090"  
 /chromosome="10"  
 /cell\_type="ES"  
 1. .1659  
 /gene="Aire"  
 1. .1659  
 /gene="Aire"  
 /codon\_start=1  
 /product="Aire protein"  
 /protein\_id="CAB36909.1"  
 /db\_xref="GI:4456675"  
 /db\_xref="GOA:O920E3"  
 /db\_xref="Swiss-Prot:O920E3"  
 /translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHNLADHDVPEDKF  
 QETLRLKEEGCPQAFHALLSWLTRDSGLDFWRLFKDYNLERYSLHSILGFP  
 KDVDLNSRKGRKPLAGPKAAVLPRPPTRKALREPRATPLASKVSFSGSLK  
 TKPKPKDGNLESQHLPLGNGIOTMAASVQRAVTVASGDPVGTGAVGILLIOVRES  
 GSKKCIQVGGFPTPNKEDPSGNLKKKASGSLKPYRAKAGQVITPGRDEQVQ  
 OCGVPPPLPSPEPVONKNEDECAVCHDGSELICDGCPRAFHLACTLSPLOEIPS  
 GIMRSCCTIQGRVQNLISQPEVSRPELPAETPIIVGLRSASEKTRGSRBELKASDA  
 AVTVNLAPHPAPALPESALCPILSAGNEGRPGAPARSVCGDGTEVLRCACHCA  
 AAFHWRCHPPTAARPGTNLRCKSCSADSTPTGPGEAIVPTSGPRPAPGLAKVGDD  
 ASHDPVLRHLDLESLNLNHSFDGILLQMAIQMSRPLAETPPFS"

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1659;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACCTGCGGCTGCCACTTCCC 1515  
 ||||||||||||||||||||||||||||||||||||  
 DB 1348 GCCGCTGCTTCACCTGCGGCTGCCACTTCCC 1379

RESULT 24  
 AF128121 1744 bp mRNA linear ROD 29-FEB-2000  
 LOCUS AF128121  
 DEFINITION Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
 spliced product 2d, complete cds.  
 ACCESSION AF128121  
 VERSION AF128121.1 GI:7108543  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 1744)  
 Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.

TITLE Expression and alternative splicing of the mouse autoimmune regulator gene (Aire)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1744)  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Direct Submission  
JOURNAL Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW Archer Road, Gainesville, FL 32610, USA  
FEATURES  
source 1..1744  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1744  
/gene="Aire"  
58..1527  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 2d"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF3646.1"  
/db\_xref="GI:7108544"  
/translation="MAGDGMRLRLRLHRTETIAVAIDASAPPLHLADHDVPEDEK  
DETLRKEGCGPOAFHALLSWLITRDSGAILDFWILPKDYNLEYSRLHSILDGFP  
KVDLNSRKGRKPLAGKAVLPPTPKRRLAEPRATPATLASKSVSGSHLK  
TKPKKPDGNLESQHLPLANGIQTMASVQRAVTVASGVPEGRVAGETILIQVPE  
GSKKCIQVGGSEFTYFNKPEDPSGNLKNKASGSLKPVARKAGQDEQVGGQCG  
VPLPLPSEPDQVQNLQPEVSRPELPATPEAPGAPARCSVCGDTEVLRCAHCAAFHW  
SCCLQGVQNLQPEVSRPELPATPEAPGAPARCSVCGDTEVLRCAHCAAFHW  
CHPPTAARPGTNLRCKSCSDSTPTPTGTPGPAVPSGPRAPGLAKVDDSDASHPV  
LHRDLSLNLNHSFDGILQWAIQSMRPLAETPPFS"

ORIGIN  
Query Match 1.6%; Score 32; DB 10; Length 1744;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACCTGCGCGCCACCTTCCC 1515  
|||||  
1213 GCGCTGCTTCACCTGCGCGCCACCTTCCC 1244

RESULT 25  
AF128120 1747 bp mRNA linear ROD 29-FEB-2000  
LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 2c, complete cds.  
ACCESSION AF128120  
VERSION AF128120.1 GI:7108541  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1747)  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Expression and alternative splicing of the mouse autoimmune regulator gene (Aire)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1747)  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Direct Submission  
JOURNAL Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW Archer Road, Gainesville, FL 32610, USA  
FEATURES  
source 1..1747  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"

/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1747  
/gene="Aire"  
58..1527  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 2c"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF3646.1"  
/db\_xref="GI:7108542"  
/translation="MAGDGMRLRLRLHRTETIAVAIDASAPPLHLADHDVPEDEK  
DETLRKEGCGPOAFHALLSWLITRDSGAILDFWILPKDYNLEYSRLHSILDGFP  
KVDLNSRKGRKPLAGKAVLPPTPKRRLAEPRATPATLASKSVSGSHLK  
TKPKKPDGNLESQHLPLANGIQTMASVQRAVTVASGVPEGRVAGETILIQVPE  
GSKKCIQVGGSEFTYFNKPEDPSGNLKNKASGSLKPVARKAGQDEQVGGQCG  
VPLPLPSEPDQVQNLQPEVSRPELPATPEAPGAPARCSVCGDTEVLRCAHCAAFHW  
SCCLQGVQNLQPEVSRPELPATPEAPGAPARCSVCGDTEVLRCAHCAAFHW  
CHPPTAARPGTNLRCKSCSDSTPTPTGTPGPAVPSGPRAPGLAKVDDSDASHPV  
LHRDLSLNLNHSFDGILQWAIQSMRPLAETPPFS"

ORIGIN  
Query Match 1.6%; Score 32; DB 10; Length 1747;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACCTGCGCGCCACCTTCCC 1515  
|||||  
1216 GCGCTGCTTCACCTGCGCGCCACCTTCCC 1247

RESULT 26  
AF128119 1756 bp mRNA linear ROD 29-FEB-2000  
LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 2b, complete cds.  
ACCESSION AF128119  
VERSION AF128119.1 GI:7108539  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1756)  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Expression and alternative splicing of the mouse autoimmune regulator gene (Aire)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1756)  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Direct Submission  
JOURNAL Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW Archer Road, Gainesville, FL 32610, USA  
FEATURES  
source 1..1756  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1756  
/gene="Aire"  
58..1536  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 2b"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF3646.1"  
/db\_xref="GI:7108540"

/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHALADHDVPEDKF  
OETLRKEKEGCPQAFHALSWLITRDSALIDFWRILFKDYNLERYSLHSLIDGFP  
KVDVLDNQRKGRKPLAGPRAAVLPPTKRALBEPATPATLASKSVSPGSHLK  
TKPPKPDNLESQHLPLGNGIQTMAASVQRAVTVASGVPGTRGAVEGILLIQVFE  
GSKKCIQVGEFYTYPNKFEDPSGNLKNRAGSSSLKPVRAKAGVITLPGDEOKV  
OOCGVPLPLSPBPVQNKDECAVCHDGSELIICDGPRAFLACLSPLOEIPIS  
LWRSCCLQGRVQONLSQPEVSRPELPATPPTGPAVSARCSVCGDTEVLRCACHCA  
AFHWRCHFTPAARPGTNLRKCSADSTPTPTGPAVSARCSVCGDTEVLRCACHCA  
HDPVLRDDESLNHSFDGILQMAIQMSRPLAETPPFS"

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1756;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1484 GCCGCTGCTTCACACTGGCGCTGCCACTTCCC 1515  
Db 1225 GCCGCTGCTTCACACTGGCGCTGCCACTTCCC 1256

RESULT 27  
AF128118 1759 bp mRNA linear ROD 29-FEB-2000  
LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 2a, complete cds.  
AF128118  
ACCESSION AF128118.1 GI:7108537  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM

REFERENCE  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Expression and alternative splicing of the mouse autoimmune  
regulator gene (Aire)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1759)  
AUTHORS Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
TITLE Direct Submission  
JOURNAL Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
Archer Road, Gainesville, FL 32610, USA

FEATURES  
source location/Qualifiers  
1..1759

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1759  
/gene="Aire"  
58..1539  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 2a"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF36463.1"  
/db\_xref="GI:7108538"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHALADHDVPEDKF  
OETLRKEKEGCPQAFHALSWLITRDSALIDFWRILFKDYNLERYSLHSLIDGFP  
KVDVLDNQRKGRKPLAGPRAAVLPPTKRALBEPATPATLASKSVSPGSHLK  
TKPPKPDNLESQHLPLGNGIQTMAASVQRAVTVASGVPGTRGAVEGILLIQVFE  
GSKKCIQVGEFYTYPNKFEDPSGNLKNRAGSSSLKPVRAKAGVITLPGDEOKV  
OOCGVPLPLSPBPVQNKDECAVCHDGSELIICDGPRAFLACLSPLOEIPIS  
LWRSCCLQGRVQONLSQPEVSRPELPATPPTGPAVSARCSVCGDTEVLRCACHCA  
AFHWRCHFTPAARPGTNLRKCSADSTPTPTGPAVSARCSVCGDTEVLRCACHCA  
HDPVLRDDESLNHSFDGILQMAIQMSRPLAETPPFS"

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1759;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1484 GCCGCTGCTTCACACTGGCGCTGCCACTTCCC 1515  
Db 1228 GCCGCTGCTTCACACTGGCGCTGCCACTTCCC 1259

RESULT 28  
MM0243821 1906 bp mRNA linear ROD 14-JAN-2000  
LOCUS Mus musculus mRNA for APECED protein (AIRE gene).  
DEFINITION AJ243821  
ACCESSION AJ243821.1 GI:6706792  
VERSION  
KEYWORDS Aire gene; APECED protein.  
SOURCE Mus musculus (house mouse)  
ORGANISM

REFERENCE  
AUTHORS Halonen,M., Pelto-Huikko,M., Palvimo,J., Björnsen,P., Peltonen,L.,  
Uimari,M., and Kolmer,M.  
TITLE Expression of the mouse Aire  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1906)  
AUTHORS Kolmer,M.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-1999) Kolmer M., Department of Human Molecular  
Genetics, National Public Health Institute, Mannerheimintie 166,  
Helsinki, FIN-00300, FINLAND

FEATURES  
source location/Qualifiers

1..1906  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..1906  
/gene="AIRE"  
1..1659  
/gene="AIRE"  
/function="putative role in transcriptional activation"  
/codon\_start=1  
/product="APECED protein"  
/protein\_id="CAB6141.1"  
/db\_xref="GI:6706793"  
/db\_xref="GOA:O920E3"  
/db\_xref="Swiss-Prot:O920E3"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHALADHDVPEDKF  
OETLRKEKEGCPQAFHALSWLITRDSALIDFWRILFKDYNLERYSLHSLIDGFP  
KVDVLDNQRKGRKPLAGPRAAVLPPTKRALBEPATPATLASKSVSPGSHLK  
TKPPKPDNLESQHLPLGNGIQTMAASVQRAVTVASGVPGTRGAVEGILLIQVFE  
GSKKCIQVGEFYTYPNKFEDPSGNLKNRAGSSSLKPVRAKAGVITLPGDEOKV  
OOCGVPLPLSPBPVQNKDECAVCHDGSELIICDGPRAFLACLSPLOEIPIS  
LWRSCCLQGRVQONLSQPEVSRPELPATPPTGPAVSARCSVCGDTEVLRCACHCA  
AFHWRCHFTPAARPGTNLRKCSADSTPTPTGPAVSARCSVCGDTEVLRCACHCA  
HDPVLRDDESLNHSFDGILQMAIQMSRPLAETPPFS"

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1906;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1484 GCCGCTGCTTCACACTGGCGCTGCCACTTCCC 1515  
Db 1348 GCCGCTGCTTCACACTGGCGCTGCCACTTCCC 1379

RESULT 29  
AF128117 1921 bp mRNA linear ROD 29-FEB-2000  
LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 1d, complete cds.  
AF128117  
ACCESSION AF128117.1 GI:7108535

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1921)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Expression and alternative splicing of the mouse autoimmune  
regulator gene (Aire)  
Unpublished  
2 (bases 1 to 1921)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Direct Submission  
Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
Archer Road, Gainesville, FL 32610, USA  
Location/Qualifiers  
1..1921  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1921  
/gene="Aire"  
58..1701  
/function="transcription factor"  
/note="alternatively spliced product id"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF36462.1"  
/db\_xref="GI:7108536"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHALADHDVPEKDF  
OPTLRKKEGCPQAFHALLSWLTTRDSCALIDFRLTPEKYNLEYSRLSLIDGFP  
KVDLNLQSKRGKPLAGPKAVLPPTPTKRLALEBPATTPATLASKVSRSQSHLK  
TKPPKPDGNLSQHLPLGNGIQTMASVQRAVAVASGDVPGTRGAVGILIQVFES  
GRSKKCIQVGEFYTYPNKFEDPSGNLKNKARSGSLKPVRAKGAQGRDEQVQCG  
VPLPSLPSBPQVONKNEDECAVCHDGEILICDGCPRAFILACUSPJOETPSGLMR  
CSCCIGRVQONLSQEVSRPELPATPILVGLRSASEKTRGSPRELRKASDAVATY  
VNLAPRAAPLILPSALCPILASNGNRPAPASRCSVCGDTEVLRCAHCAAAFH  
RCHFTPAARPGTNLRCKSCSADSTPTPTGTEAVPTSGPRPAFLAKVDDDSASHD  
VLRHDLLESLNHSFDGILQWAIQSMRPLAETPPSS"

ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1921;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCCTCCACTGGCGCTGCCACTTCCC 1515  
|||||  
DB 1390 GCCGCTGCTTCCTCCACTGGCGCTGCCACTTCCC 1421

RESULT 30  
AF128116 1924 bp mRNA linear ROD 29-FEB-2000  
LOCUS AF128116 Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 1c, complete cds.  
ACCESSION AF128116 GI:7108533  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1924)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Expression and alternative splicing of the mouse autoimmune  
regulator gene (Aire)  
Unpublished  
2 (bases 1 to 1924)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Direct Submission  
Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
Archer Road, Gainesville, FL 32610, USA  
Location/Qualifiers  
1..1924  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1924  
/gene="Aire"  
58..1704  
/function="transcription factor"  
/note="alternatively spliced product 1c"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF36461.1"  
/db\_xref="GI:7108534"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHALADHDVPEKDF  
OPTLRKKEGCPQAFHALLSWLTTRDSCALIDFRLTPEKYNLEYSRLSLIDGFP  
KVDLNLQSKRGKPLAGPKAVLPPTPTKRLALEBPATTPATLASKVSRSQSHLK  
TKPPKPDGNLSQHLPLGNGIQTMASVQRAVAVASGDVPGTRGAVGILIQVFES  
GRSKKCIQVGEFYTYPNKFEDPSGNLKNKARSGSLKPVRAKGAQGRDEQVQCG  
VPLPSLPSBPQVONKNEDECAVCHDGEILICDGCPRAFILACUSPJOETPSGLMR  
CSCCIGRVQONLSQEVSRPELPATPILVGLRSASEKTRGSPRELRKASDAVATY  
VNLAPRAAPLILPSALCPILASNGNRPAPASRCSVCGDTEVLRCAHCAAAFH  
RCHFTPAARPGTNLRCKSCSADSTPTPTGTEAVPTSGPRPAFLAKVDDDSASHD  
VLRHDLLESLNHSFDGILQWAIQSMRPLAETPPSS"

ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1924;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCCTCCACTGGCGCTGCCACTTCCC 1515  
|||||  
DB 1393 GCCGCTGCTTCCTCCACTGGCGCTGCCACTTCCC 1424

RESULT 31  
AF128115 1933 bp mRNA linear ROD 29-FEB-2000  
LOCUS AF128115 Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 1b, complete cds.  
ACCESSION AF128115 GI:7108531  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1933)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Expression and alternative splicing of the mouse autoimmune  
regulator gene (Aire)  
Unpublished  
2 (bases 1 to 1933)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Direct Submission  
Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
Archer Road, Gainesville, FL 32610, USA  
Location/Qualifiers  
1..1933  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1933

## CDS

/gene="Aire"  
58..1713  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 1b"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAP16460.1"  
/db\_xref="GI:7108532"

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1933;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACATGCGGCGCCACTTCCC 1515  
|||||  
Db 1402 GCCGCTGCTTCACATGCGGCGCCACTTCCC 1433

RESULT 32  
LOCUS AF079536 1936 bp mRNA linear ROD 17-MAR-1999  
DEFINITION Mus musculus autoimmune regulator (Aire) mRNA, complete cds.  
ACCESSION AF079536  
VERSION AF079536.1 GI:4426598  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1936)  
Wang,C.Y., Shi,J.D., Davoodi-Semlomi,A. and She,J.X.  
TITLE Cloning of Aire, the mouse homologue of the autoimmune regulator (AIRE) gene responsible for autoimmune polyglandular syndrome type 1 (ASPI)

JOURNAL Genomics 55 (3), 322-326 (1999)

MEDLINE 99168902

PUBMED 10049587

REFERENCE 2 (bases 1 to 1936)

Wang,C.Y., Shi,J.D., Davoodi-Semlomi,A. and She,J.X.  
TITLE Direct Submision  
SUBMITTED (21-JUN-1998) Pathology, Immunology and Laboratory  
Medicine, University of Florida, 1600 SW Archer Road, Gainesville,  
FL 32610, USA

FEATURES  
source location/Qualifiers

1..1936  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/cfeature\_type="thymus"  
1..1936  
/gene="Aire"  
58..1716  
/gene="Aire"  
/function="transcription factor"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAD2044.1"  
/db\_xref="GI:4426598"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHLALADHDVPEDKF  
QETRLKEKEGCPQAFHALLSWLTLTRDGAALDPFRILFKQVNLERSYSLHSLDGGP

## CDS

## ORIGIN

Query Match 1.6%; Score 32; DB 10; Length 1936;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACATGCGGCGCCACTTCCC 1515  
|||||  
Db 1405 GCCGCTGCTTCACATGCGGCGCCACTTCCC 1436

RESULT 33  
LOCUS AF105002 18351 bp DNA linear ROD 26-OCT-2001  
DEFINITION Mus musculus autoimmune regulator (Aire) gene, complete cds.  
ACCESSION AF105002  
VERSION AF105002.1 GI:5669675  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 18351)  
Shi,J.D., Wang,C.Y., Marron,M.P., Ruan,Q.Q., Huang,Y.Q.,  
Detter,J.C. and She,J.X.  
TITLE Chromosomal localization and complete genomic sequence of the  
murine autoimmune regulator gene (Aire)

JOURNAL Autoimmunity 31 (1), 47-53 (1999)

MEDLINE 20059142

PUBMED 10593569

REFERENCE 2 (bases 1 to 18351)

Shi,J.D., Wang,C.Y., Marron,M.P., Ruan,Q.Q., Huang,Y.Q.,  
Detter,J.C. and She,J.X.  
TITLE Direct Submision

SUBMITTED (06-NOV-1998) Department of Pathology, Immunology and  
Laboratory Medicine, University of Florida, 1600 SW Archer Rd.,  
Room D6-15, Gainesville, FL 32610, USA

FEATURES  
source location/Qualifiers

1..18351  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="10"  
<4918..>18138  
/gene="Aire"  
join(<4918..5052,5365..5539,5769..5925,6219..6292,  
6740..6853,8274..8422,8611..8691,10373..10488,  
11577..11676,12530..12706,13659..13780,13879..13987,  
16489..16557,18058..>18138)  
/product="autoimmune regulator"  
/gene="Aire"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAD4642.1"

## CDS

1..18351  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="10"  
<4918..>18138  
/gene="Aire"  
join(<4918..5052,5365..5539,5769..5925,6219..6292,  
6740..6853,8274..8422,8611..8691,10373..10488,  
11577..11676,12530..12706,13659..13780,13879..13987,  
16489..16557,18058..>18138)  
/product="autoimmune regulator"  
/gene="Aire"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAD4642.1"  
/db\_xref="GI:5669676"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHLALADHDVPEDKF  
QETRLKEKEGCPQAFHALLSWLTLTRDGAALDPFRILFKQVNLERSYSLHSLDGGP  
TKPKKPGNLSOHLPLNGIOTMAASVQAVTVASGDPGTRGAVEGILLQOVES  
GRSKKCIQVGEFTTPNKFEDPSGNLKNKASGSLKRVNAKAGQVTTIPRDEKVG

## ORIGIN

OQCGVPLPSLPSEPOVNOXNDECAVCHDGEELICDGCPRAPHLACLSPPLQIIPS  
GLMRCSCCLQGRVQONLSQPEVSRPELPAPETPIIVGARSASEKTRGPERELKASDA  
AVTYVNLAPHPAPALPBPALCPPLSGNBRPGPAPARSVCVGDGTEVLRCAHCA  
AAFWRCHEPTTAARPGTNLRCKSCSADSTPTGCGEAVPTSGRPRAGLAKVGDSDS  
ASHDVLHRDLESILNHSFDGIIQWAIQMSRPLAETPPSS"

Query Match 1.6%; Score 32; DB 10; Length 18351;  
Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCCTTCACCTGCGCTGCCACTTCCC 1515  
|||||  
Db 13728 GCCGCTGCCTTCACCTGCGCTGCCACTTCCC 13759

RESULT 34  
MMU007715 18616 bp DNA linear ROD 11-MAR-1999

LOCUS MMU007715 Mus musculus Aire gene.

DEFINITION AJ007715

ACCESSION AJ007715.1 GI:3550507

KEYWORDS Aire gene; autoimmune regulator.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Chordata; Mammalia; Euteleostomi;  
Eukaryota; Euteleostomi; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euteleostomi; Chordata; Vertebrata; Muridae; Mus.

REFERENCE 1 Miltaz, L., Roesier, C., Heino, M., Peterson, P., Kohn, K.J., Gos, A.,  
Morris, M.A., Kudoh, J., Shimizu, N., Antonarakis, S.E. and Scott, H.S.

JOURNAL Isolation and characterization of the mouse Aire gene

MEDLINE Biochem. Biophys. Res. Commun. 255 (2), 483-490 (1999)

PUBMED 99160890

REFERENCE 2 (bases 1 to 18616)

AUTHORS Miltaz, L.

TITLE Direct Submission

JOURNAL Submitted (26-AUG-1998) Division of Medical Genetics, University of

Geneva, 1, Rue Michel-Servet, Geneva, GE 1211, Switzerland

LOCATION/Qualifiers

FEATURES

source

1.18616

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="129Sv"

/db\_xref="taxon:10090"

complement(join(1347..2537,5851..6020,6085..6232,

6233..6296,7758..7893,10146..10356,10359..10551,

11453..11562,12262..12408,12415..12431,12432..12626,

12627..12769,12813..12949,13467..13528,13990..14210,

15304..15453,17779..17950,18535..18616))

/gene="Aire"

complement(1347..1546)

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

repeat\_unit

mrna

cds

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

/gene="Aire"  
join(2970..3104,3417..3591,3821..3976,4358..4432,  
4480..4993,6414..6562,6751..6831,8506..8621,9710..9809,  
10661..10837,11787..11908,12005..12113,14605..14673,  
16166..16246)

/gene="Aire"  
join(2970..3104,3417..3591,3821..3976,4358..4432,  
4480..4993,6414..6562,6751..6831,8506..8621,9710..9809,  
10661..10837,11787..11908,12005..12113,14605..14673,  
16166..16246)

/gene="Aire"

/codon\_start=1

/product="autoimmune regulator"

/protein\_id="CAA07620.1"

/db\_xref="GI:3550508"

/db\_xref="GOA:O920E3"

/db\_xref="Swiss-Prot:O920E3"

/translation="MAGDGMRLRLRLHRTIEIAVAIDAPPLHLADHDVPEDKF  
QETLRLKEKGGPPQAFHALLSMLTRDGAALDIPWRILFKYNLERYSLHSLDGF  
KDVLNQSRKSRPLAGPKAAVLPRPTKRLAEPPATPATLASKSVSPGSHLK  
TKPKKPPGNLESQHLPLGNGIOTMAASVORAVTVAGDVPTRGAVGIILOVFE  
GRSKKCIQVGEFTYFNPKEDPSGNLKKKASGSLKRVAVRAGKQVITPGDEQKVG  
QOCGVPLPSLPSEPOVNOXNDECAVCHDGEELICDGCPRAPHLACLSPPLQIIPS  
GLMRCSCCLQGRVQONLSQPEVSRPELPAPETPIIVGARSASEKTRGPERELKASDA  
AVTYVNLAPHPAPALPBPALCPPLSGNBRPGPAPARSVCVGDGTEVLRCAHCA  
AAFWRCHEPTTAARPGTNLRCKSCSADSTPTGCGEAVPTSGRPRAGLAKVGDSDS  
ASHDVLHRDLESILNHSFDGIIQWAIQMSRPLAETPPSS"

<2970..3104

/gene="Aire"

/number=1

3015..3416

/gene="Aire"

/number=1

3417..3591

/gene="Aire"

/number=2

3592..3820

/gene="Aire"

/number=2

3821..3976

/gene="Aire"

/number=3

3977..4357

/gene="Aire"

/number=3

4358..4432

/gene="Aire"

/number=4

4433..4879

/gene="Aire"

/number=4

4880..4993

/gene="Aire"

/number=5

4994..6413

/gene="Aire"

/number=5

6414..6562

/gene="Aire"

/number=6

6563..6750

/gene="Aire"

/number=6

6751..6831

/gene="Aire"

/number=6

6832..6980

/gene="Aire"

/number=6

6981..7129

/gene="Aire"

/number=6

7130..7278

/gene="Aire"

/number=6

7279..7427

/gene="Aire"

/number=6

7428..7576

/gene="Aire"

/number=6

7577..7725

/gene="Aire"

/number=6

7726..7874

/gene="Aire"

/number=6

7875..8023

/gene="Aire"

/number=6

8024..8172

/gene="Aire"

/number=6

8173..8321

/gene="Aire"

/number=6

8322..8470

```

/gene="Aire"
/number=7
6832..8505
/gene="Aire"
/number=7
7002..7159
/gene="Aire"
/rpt_family="SINE/B4"
7027..7160
/gene="Aire"
/rpt_family="SINE/Alu"
complement(7758..7893)
/gene="Aire"
/rpt_family="SINE/B2"
8506..8621
/gene="Aire"
/number=8
8622..9709
/gene="Aire"
/number=8
9312..9458
/gene="Aire"
/rpt_family="SINE/Alu"
9710..9809
/gene="Aire"
/number=9
9810..10660
/gene="Aire"
/number=9
complement(10146..10356)
/gene="Aire"
/rpt_family="SINE/B2"
complement(10359..10551)
/gene="Aire"
/rpt_family="SINE/B2"
10661..10837
/gene="Aire"
/number=10
10838..11786
/gene="Aire"
/number=10
11060..11182
/gene="Aire"
/rpt_family="SINE/B4"
complement(11453..11562)
/gene="Aire"
/rpt_family="LINE/L1"
11787..11908
/gene="Aire"
/number=11
11909..12004
/gene="Aire"
/number=11
12005..12113
/gene="Aire"
/number=12
12114..14604
/gene="Aire"
/number=12
complement(12262..12378)
/gene="Aire"
/rpt_family="SINE/Alu"
complement(12277..12408)
/gene="Aire"
/rpt_family="LTR/MALR"
complement(12415..12431)
/gene="Aire"
/rpt_family="SINE/B2"
complement(12432..12626)
/gene="Aire"
/rpt_family="SINE/B2"
complement(12627..12769)

```

```

Query Match      1.6%; Score 32; DB 10; Length 18616;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1484 GCCGCTGCGCTTCACCTGCGGCTGCACCTCC 1515
          |||||
DB      11856 GCCGCTGCGCTTCACCTGCGGCTGCACCTCC 11887

RESULT 35
AF073797
LOCUS
DEFINITION
AF073797
ACCESSION
AF073797
VERSION
AF073797.1 GI:4091972
KEYWORDS
HTG.
SOURCE
MUS
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 46872)
Aaltonen,J., Björnsen,P., Perheentupa,J., Horelli-Kuitunen,N.,
Palotie,A., Peltonen,L., Lee,Y.S., Francis,F., Hennig,S., Thiel,C.,
Lehrach,H. and Yaspo,M.-L.
An autoimmune disease, APECED, caused by mutations in a novel gene
featuring two PHD-type zinc-finger domains. The Finnish-German
APECED Consortium.
Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy
Nat. Genet. 17 (4), 399-403 (1997)
98061087
JOURNAL
MEDLINE
PUBMED
9398840
2 (bases 1 to 46872)
Bleeschmidt,K., Schweiger,M., Wertz,K., Poulsen,R.,
Christensen,H.M., Rosenthal,A., Lehrich,H. and Yaspo,M.L.
The mouse Aire gene: comparative genomic sequencing, gene
organization, and expression
Genome Res. 9 (2), 158-166 (1999)
10022980
JOURNAL
MEDLINE
PUBMED
99148139
3 (bases 1 to 46872)
Bleeschmidt,K., Huang-M.Y.C., Nordstiek,G., Drescher,B.,
Rosenthal,A. and Yaspo,M.-L.
Direct Submission
Submitted (22-JUN-1998) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
Location/Qualifiers
1..46872
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Ola"
/db_xref="taxon:10090"
/note="Cosmid MPMGcl2112287 originating from cosmid
library number 121 (RZPD-berlin). Cloning vector Lawriest
7."
complement(156..327)
/note="Genscan, score = 14.95%, comment = Internal_exon
172 bp frame: 2 phase: 1"
/evidence=not experimental
complement(156..327)
/note="GRAIL, score = 99.000%, comment = excellent"
/evidence=not experimental
complement(156..316)
/note="Xpound exon prediction, score = 93% (0%)"
/evidence=not experimental
complement(935..1047)
/note="Genscan, score = 11.98%, comment = Internal_exon
113 bp frame: 0 phase: 2"
/evidence=not experimental
complement(935..1047)
/note="MZER, score = 96.5%"
/evidence=not experimental
complement(935..1047)

```



```

/note="GRAIL, score = 100.000%, comment = excellent"
/evidence=not experimental
complement(1129..1208)
/note="GenScan, score = 3.57%, comment = Internal_exon 80
bp frame: 0 phase: 2"
/evidence=not experimental
complement(1129..1208)
/note="MZF, score = 57%"
/evidence=not experimental
complement(1572..1634)
/evidence=not experimental
/rpt_family="AluSp"
/rpt_unit="gga99g"
/note="MZF, score = 91.3%"
/evidence=not experimental
complement(1805..1885)
/evidence=not experimental
/rpt_family="AluJ/FLAM"
complement(2079..2318)
/note="GenScan, score = 4.62%, comment = Internal_exon 240
bp frame: 0 phase: 0"
/evidence=not experimental
complement(2079..2241)
/note="MZF, score = 79.6%"
/evidence=not experimental
complement(2079..2123)
/note="GRAIL, score = 99.000%, comment = excellent"
/evidence=not experimental
complement(2079..2121)
/note="Xpound exon prediction, score = 84% (0%) "
/evidence=not experimental
complement(2170..2318)
/note="GRAIL, score = 64.000%, comment = good"
/evidence=not experimental
complement(2355..2441)
/note="GenScan, score = 5.24%, comment = Internal_exon 87
bp frame: 0 phase: 0"
/evidence=not experimental
complement(2355..2441)
/note="MZF, score = 99.1%"
/evidence=not experimental
complement(2355..2441)
/note="GRAIL, score = 88.000%, comment = excellent"
/evidence=not experimental
complement(2922..3023)
/note="GRAIL, score = 70.000%, comment = good"
/evidence=not experimental
complement(3128..3173)
/note="GenScan, score = 3.08%, comment = Internal_exon 46
bp frame: 2 phase: 1"
/evidence=not experimental
complement(3128..3173)
/note="GRAIL, score = 94.000%, comment = excellent"
/evidence=not experimental
complement(3128..3163)
/note="MZF, score = 75.9%"
/evidence=not experimental
3243..3410
/note="homology = 59.50%, score = 26, counts = 28"
/evidence=not experimental
/rpt_type=tandem
/rpt_unit="gga99g"
complement(3847..3975)
/note="GenScan, score = 1.75%, comment = Initial_exon 129
bp frame: 0 phase: 0"
/evidence=not experimental
complement(3847..3975)
/note="GRAIL, score = 56.000%, comment = good"
/evidence=not experimental
4343..4367
/standard_name="gap"
/note="sequence gap of unknown size"
complement(4559..4630)

```

```

/evidence=not experimental
/rpt_family="AluSg/x"
complement(4757..4984)
/evidence=not experimental
/rpt_family="AluOb"
complement(7733..7878)
/note="GRAIL, score = 57.000%, comment = good shadow"
/evidence=not experimental
7794..7905
/note="GRAIL, score = 52.000%, comment = good"
/evidence=not experimental
complement(8914..9038)
/evidence=not experimental
/rpt_family="AluSg"
9395..9666
/note="CpG island score = 0.74, GC = 63.20%, CpGs = 25;
Region: CpG island"
/evidence=not experimental
9555..22828
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
9555..9689
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
/number=1
/evidence=experimental
10002..10176
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
/number=2
/evidence=experimental
10406..10561
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
/number=3
/evidence=experimental
10943..11017
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
/number=4
/evidence=experimental
11465..11578
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
/number=5
/evidence=experimental
complement(12241..12306)
/evidence=not experimental
/rpt_family="MUTID"
complement(12539..12600)
/evidence=not experimental
/rpt_family="tRNA-Cys-TGC"
complement(12664..12819)
/evidence=not experimental
/rpt_family="AluSg"
complement(12820..12881)
/evidence=not experimental
/rpt_family="MUTIC"
12999..13147
/gene="Mouse homologue of the human APECD gene (AIRE;
acc. No. Z97990) "
/number=6
/evidence=experimental
13336..13416
/gene="Mouse homologue of the human APECD gene (AIRE;

```

Query Match 1.6%; Score 32; DB 10; Length 46872;  
Best Local Similarity 100.0%; Pred. No. 2.8e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1484 GCCGCTGCCCTTCACCTGCGCGTCCCACTTCCC 1515
DB 18441 GCCGCTGCCCTTCACCTGCGCGTCCCACTTCCC 18472

```

```

RESULT 36
AC015891
LOCUS
DEFINITION
AC015891 158049 bp DNA 1linear HTG 20-NOV-2002
MUS MUSCULUS CHROMOSOME 10 CLONE RP21-522L13 MAP 10, *** SEQUENCING
IN PROGRESS *** 6 UNORDERED PIECES.
AC015891
VERSION
AC015891.16 GI:25140119
KEYWORDS
HTG: HTGS_PHASE1; HTGS_FULTOP; HTGS_ACTIVEFLN.
SOURCE
MUS MUSCULUS (HOUSE MOUSE)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 158049)
Birren, B., Nusbaum, C. and Lander, E.
MUS MUSCULUS CHROMOSOME 10, CLONE RP21-522L13
Unpublished
2 (bases 1 to 158049)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baltwin, J., Bartha, N., Beckerly, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hages, B., Hatford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, D.,
Lehoczky, J., Lien, C., Locke, K., MacDonald, P., Margulis, N.,
McEwan, P., McGuirk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Teefaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, M. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 158049)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Bartha, N., Baetjen, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camata, U., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrim, J., Menues, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Teefaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Willson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (20-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 20, 2002 this sequence version replaced gi:25046447.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L761
Center clone name: 522_L_13
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces

```

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 58675: contig of 58675 bp in length
* 58676 58775: gap of 100 bp
* 58776 75633: contig of 16858 bp in length
* 75634 75733: gap of 100 bp
* 75734 78147: contig of 2414 bp in length
* 78148 78247: gap of 100 bp
* 78248 107125: contig of 28878 bp in length
* 107126 107225: gap of 100 bp
* 107226 143742: contig of 36517 bp in length
* 143743 143842: gap of 100 bp
* 143843 158049: contig of 14207 bp in length.
Location/Qualifiers
1. 158049
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10"
/map="10"
/clone="RP21-522L13"
/clone_lib="RP21-522L13 Female Mouse PAC"

ORIGIN
Query Match 1.6%; Score 32; DB 2; Length 158049;
Best Local Similarity 100%; Pred. No. 2.4e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACTGCGCGCCACTTCCC 1515
DB 156239 GCCGCTGCTTCACTGCGCGCCACTTCCC 156260

RESULT 37
AC138672/c
LOCUS
DEFINITION
AC138672 190019 bp DNA 1linear HTG 22-MAR-2003
MUS MUSCULUS CHROMOSOME 10 CLONE RP23-411J14 MAP 10, WORKING DRAFT
SEQUENCE, 13 UNORDERED PIECES.
AC138672
VERSION
AC138672.3 GI:29150492
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
MUS MUSCULUS (HOUSE MOUSE)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 190019)
Birren, B., Nusbaum, C. and Lander, E.
MUS MUSCULUS CHROMOSOME 10, CLONE RP23-411J14
Unpublished
2 (bases 1 to 190019)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Bartha, N., Baetjen, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camata, U., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrim, J., Menues, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Teefaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Willson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission

```



TITLE Mus musculus SNPs  
JOURNAL Unpublished (2003)  
COMMENT

Contact: Jonathan Ueoka  
Roche Palo Alto Genetics and Genomics Department  
Roche Palo Alto  
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA  
Tel: 6508555807  
Email: Jonathan.Ueoka@roche.com  
Primer A: No primer submitted.  
Primer B: No primer submitted.

FEATURES  
source  
1..496  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/map="10-3106-3535-CANA01111727.1.1.11206"  
/clone\_1b="Roche Palo Alto"  
/note="SNPs developed from assay sequences derived from 15 different strains-of mice (as of October 1, 2003). Those strains include A/J, A/HeJ, -129/Sv, AKR/J, B10.D2-H2/GSnJ, BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, -CAST/Ei, DBA/2J, MRL/MPJ, NZB/BinJ, NZW/Lac, SPRET/Ei.-"

ORIGIN  
STS  
Query Match 1.5%; Score 30; DB 11; Length 496;  
Best Local Similarity 100.0%; Pred. No. 0.00067;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1715 GGCACTCTGCACTGGCCATCCAGAGCATG 1744  
|||||  
Db 136 GGCACTCTGCACTGGCCATCCAGAGCATG 107

RESULT 39  
AF128125 1622 bp mRNA linear ROD 29-FEB-2000  
LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 3d, complete cds.  
ACCESSION AF128125  
VERSION AF128125.1 GI:7108551  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 1622)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Expression and alternative splicing of the mouse autoimmune  
regulator gene (Aire)  
Unpublished  
JOURNAL  
REFERENCE  
2 (bases 1 to 1622)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Direct Submision  
TITLE Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
JOURNAL Archer Road, Gainesville, FL 32610, USA  
AUTHORS  
Archer Road, Gainesville, FL 32610, USA  
FEATURES  
source  
1..1622  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1622  
/gene="Aire"  
58..1272  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 3d"  
/codon\_start=1  
/product="autoimmune regulator"

/protein\_id="AA36470.1"  
/db\_xref="GI:7108552"  
/translation="MAGDGMRLRLRLHRTETIAVAIDSAFPLHALADHVPEDEK  
QETTLKKEKCCPQAFHALLSWLTTRDSGALIDFWRLIFKDYNLERYRLSHLILGFP  
KDVDLNSRRKRPKPLAGPKAAVLPPRPPTKRLALEBPATPATLASKVSFSGHLK  
TKPDKPKDGNLESCHLPLGNGIQTMAASVORAIVAGDVAGTGAVERGILIQVFES  
GRSKKCIQVGESEFTYFNKEDPDSGLKKKASGSLKRVVAKAGQGRBOGVGQCC  
VPPLPSLSEPOVONNEDECAVCHDGDELICCGCPRAFLACISPLQETPSGLMR  
SCCLQGRVQONLSQPEVSRPELPATRPDPSPLQIILCRDLSHARHTGRSCTHLMAPS  
STWACQGRRLC"

ORIGIN  
Query Match 1.5%; Score 30; DB 10; Length 1622;  
Best Local Similarity 100.0%; Pred. No. 0.00056;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1715 GGCACTCTGCACTGGCCATCCAGAGCATG 1744  
|||||  
Db 1334 GGCACTCTGCACTGGCCATCCAGAGCATG 1363

RESULT 40  
AF128124 1625 bp mRNA linear ROD 29-FEB-2000  
LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
DEFINITION spliced product 3c, complete cds.  
ACCESSION AF128124  
VERSION AF128124.1 GI:7108549  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 1625)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Expression and alternative splicing of the mouse autoimmune  
regulator gene (Aire)  
Unpublished  
JOURNAL  
REFERENCE  
2 (bases 1 to 1625)  
Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
Direct Submision  
TITLE Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
JOURNAL Archer Road, Gainesville, FL 32610, USA  
AUTHORS  
Archer Road, Gainesville, FL 32610, USA  
FEATURES  
source  
1..1625  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="B6"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="between D10Mit31 and D10Mit10"  
1..1625  
/gene="Aire"  
58..1275  
/gene="Aire"  
/function="transcription factor"  
/note="alternatively spliced product 3c"  
/codon\_start=1  
/product="autoimmune regulator"  
/protein\_id="AAF36469.1"  
/db\_xref="GI:7108550"

ORIGIN  
Query Match 1.5%; Score 30; DB 10; Length 1625;  
Best Local Similarity 100.0%; Pred. No. 0.00056;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 GGCATCTGCGAGGGCCATCCAGAGCATG 1744  
 |||||  
 DB 1337 GGCATCTGCGAGGGCCATCCAGAGCATG 1366

RESULT 41  
 AF128123 1634 bp mRNA linear ROD 29-FEB-2000  
 LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
 DEFINITION spliced product 3b, complete cds.  
 AF128123  
 VERSION AF128123.1 GI:7108547  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 1634)  
 Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
 Expression and alternative splicing of the mouse autoimmune  
 TITLE regulator gene (Aire)  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1634)  
 Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
 Direct Submission  
 AUTHORS Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
 TITLE Archer Road, Gainesville, FL 32610, USA  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1.1634  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="B6"  
 /db\_xref="taxon:10090"  
 /chromosome="10"  
 /map="between D10Mit31 and D10Mit10"  
 1.1634  
 /gene="Aire"  
 58.1284  
 /gene="Aire"  
 /function="transcription factor"  
 /note="alternatively spliced product 3b"  
 /codon\_start=1  
 /product="autoimmune regulator"  
 /protein\_id="AAF36468.1"  
 /db\_xref="GI:7108548"  
 /translation="MAGSDGMLRRLRLHRTIIVAVDAISAPFLHMLADHDVPEDKF  
 QETLRKKEGCPQAFHALSWLITRDSGAILDFWITLKDNLYERYSLHSILDFP  
 KOVDLNQSRKGRPLAGPKAVLPPTPKKALEBPATTPATLASKSVSPGSHLK  
 TKPKPKDGNLSQHLPLGNGIQTMAASVQRAVTVASGVPTRGAVESGILIQVYES  
 GRSKKCIQVGEFYPNKFEDPSGNLKNKARSGSLKPVRAKQAVTIPGRDEQVG  
 OOCGVPLPSLPSPEPVQNKNEDECAVCHGGLGCLCDGCPRAFLACISPLPDI  
 LMRSCCLQGRVQONLSQPEVSRPELPATRPDQSPVQILLCRLDHSHARTGRSCTH  
 MAPSSTWACQGRRLC"

ORIGIN  
 Query Match 1.5%; Score 30; DB 10; Length 1634;  
 Best Local Similarity 100.0%; Pred. No. 0.00056;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 GGCATCTGCGAGGGCCATCCAGAGCATG 1744  
 |||||  
 DB 1346 GGCATCTGCGAGGGCCATCCAGAGCATG 1375

RESULT 42  
 AF128122 1637 bp mRNA linear ROD 29-FEB-2000  
 LOCUS Mus musculus autoimmune regulator (Aire) mRNA, alternatively  
 DEFINITION spliced product 3a, complete cds.  
 AF128122  
 VERSION AF128122.1 GI:7108545

KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 1637)  
 Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
 Expression and alternative splicing of the mouse autoimmune  
 TITLE regulator gene (Aire)  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1637)  
 Ruan,Q.G., Wang,C.Y., Shi,J.D. and She,J.X.  
 Direct Submission  
 AUTHORS Submitted (12-FEB-1999) Pathology, University of Florida, 1600 SW  
 TITLE Archer Road, Gainesville, FL 32610, USA  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1.1637  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="B6"  
 /db\_xref="taxon:10090"  
 /chromosome="10"  
 /map="between D10Mit31 and D10Mit10"  
 1.1637  
 /gene="Aire"  
 58.1287  
 /gene="Aire"  
 /function="transcription factor"  
 /note="alternatively spliced product 3a"  
 /codon\_start=1  
 /product="autoimmune regulator"  
 /protein\_id="AAF36467.1"  
 /db\_xref="GI:7108546"  
 /translation="MAGSDGMLRRLRLHRTIIVAVDAISAPFLHMLADHDVPEDKF  
 QETLRKKEGCPQAFHALSWLITRDSGAILDFWITLKDNLYERYSLHSILDFP  
 KOVDLNQSRKGRPLAGPKAVLPPTPKKALEBPATTPATLASKSVSPGSHLK  
 TKPKPKDGNLSQHLPLGNGIQTMAASVQRAVTVASGVPTRGAVESGILIQVYES  
 GRSKKCIQVGEFYPNKFEDPSGNLKNKARSGSLKPVRAKQAVTIPGRDEQVG  
 OOCGVPLPSLPSPEPVQNKNEDECAVCHGGLGCLCDGCPRAFLACISPLPDI  
 LMRSCCLQGRVQONLSQPEVSRPELPATRPDQSPVQILLCRLDHSHARTGRSCTH  
 LMAPSSTWACQGRRLC"

ORIGIN  
 Query Match 1.5%; Score 30; DB 10; Length 1637;  
 Best Local Similarity 100.0%; Pred. No. 0.00056;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 GGCATCTGCGAGGGCCATCCAGAGCATG 1744  
 |||||  
 DB 1349 GGCATCTGCGAGGGCCATCCAGAGCATG 1378

RESULT 43  
 AC007433 125105 bp DNA linear ROD 30-JAN-2001  
 LOCUS Mus musculus chromosome 10, clone RP21-340M5, complete sequence.  
 AC007433  
 VERSION AC007433.16 GI:12597916  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 125105)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 10, clone RP21-340M5  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 125105)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baker,J., Baldwin,J., Barna,N., Beckert,J., Berr,J., Brown,A.,  
 Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., DeBellano,K., Depayre,B., Devon,K., Dewar,K.,  
 Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,

Funk, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kam, L., Karcas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McSwan, P., McGuire, A., McKernan, K., McLoughlin, J., Melgrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Testaye, S., Tornella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

# TITLE JOURNAL

Submitted (29-Apr-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE AUTHORS

3 (bases 1 to 125105)  
Barr, N., Baetjen, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Batten, B., Bogen, V., Boguslavsky, L., Boukhvalov, B., Brown, A., Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karcas, A., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Melgrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phukhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupbach, R., Seaman, S., Severy, P., Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strausz, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Travers, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE JOURNAL

Submitted (30-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 30, 2001 this sequence version replaced gi:5624004.  
All repeats were identified using RepeatMasker:

## COMMENT

Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RN/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L763

Center clone name: 340\_M5

## FEATURES

source

Location/Qualifiers  
1. 125105  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="10"  
/map="10"  
/clone="RP21-340M5"  
/clone\_lib="RP21-21 Female Mouse PAC"  
complement(385..605)  
repeat\_region  
/rpt\_family="URR1B"  
1355..1453  
repeat\_region  
/rpt\_family="PB1D9"  
1459..1651  
repeat\_region  
/rpt\_family="B3A"  
complement(1699..1848)  
repeat\_region  
/rpt\_family="B1\_MM"  
1873..1955  
repeat\_region  
/rpt\_family="B2\_MM2"  
1956..1996  
repeat\_region  
/rpt\_family="TG)n"  
1997..2060  
repeat\_region

repeat\_region  
/rpt\_family="B2\_MM2"  
2234..2278  
repeat\_region  
/rpt\_family="GA)n"  
3230..3380  
repeat\_region  
/rpt\_family="B2\_MM2"  
complement(4174..4345)  
repeat\_region  
/rpt\_family="B2\_MM2"  
complement(4889..5136)  
repeat\_region  
/rpt\_family="B3"  
5279..5478  
repeat\_region  
/rpt\_family="B2\_MM2"  
5610..5136  
repeat\_region  
/rpt\_family="RMR19B"  
6135..6444  
repeat\_region  
/rpt\_family="RMR19B"  
6681..6828  
repeat\_region  
/rpt\_family="B1\_MM"  
6926..7051  
repeat\_region  
/rpt\_family="B1\_MM"  
7052..7066  
repeat\_region  
/rpt\_family="B4A"  
complement(7264..7661)  
repeat\_region  
/rpt\_family="IAPLTR1\_MM"  
complement(7666..10945)  
repeat\_region  
/rpt\_family="IAPEx-int"  
9626  
unsure  
/note="Probably CC."  
complement(10936..12124)  
repeat\_region  
/rpt\_family="IAPEx-int"  
complement(12099..12261)  
repeat\_region  
/rpt\_family="IAPEx-int"  
complement(12262..12659)  
repeat\_region  
/rpt\_family="IAPLTR1\_MM"  
complement(12677..12738)  
repeat\_region  
/rpt\_family="B2\_MM2"  
complement(12743..12869)  
repeat\_region  
/rpt\_family="B1\_MM"  
12935..12988  
repeat\_region  
/rpt\_family="CT-rich"  
12990..13017  
repeat\_region  
/rpt\_family="TG)n"  
13374..13427  
repeat\_region  
/rpt\_family="TG)n"  
13525..13593  
repeat\_region  
/rpt\_family="TTTC)n"  
14148..14459  
repeat\_region  
/rpt\_family="MTD"  
14513..14583  
repeat\_region  
/rpt\_family="B1\_MM"  
complement(14624..14825)  
repeat\_region  
/rpt\_family="B3"  
14827..14879  
repeat\_region  
/rpt\_family="TRA)n"  
complement(14882..15062)  
repeat\_region  
/rpt\_family="B2\_MM2"  
complement(15063..15224)  
repeat\_region  
/rpt\_family="B2\_MM2"  
complement(15529..15905)  
repeat\_region  
/rpt\_family="RLTR11A"  
complement(16659..16861)  
repeat\_region  
/rpt\_family="B2\_MM2"  
17055..17145  
repeat\_region  
/rpt\_family="B1F"  
17164..17422  
repeat\_region  
/rpt\_family="HAL1"  
17817..17995  
repeat\_region  
/rpt\_family="Lx7"  
complement(18009..18117)  
repeat\_region  
/rpt\_family="Lx9"  
complement(18128..18279)  
repeat\_region  
/rpt\_family="B1\_MM"  
18308..18434  
repeat\_region  
/rpt\_family="Lx7"

```

repeat_region      18435..18627
                    /rpt_family="B2_Mm2"
repeat_region      18640..18676
                    /rpt_family="(TTG)n"
repeat_region      18806..18806
                    /rpt_family="B1_MM"
repeat_region      18821..18899
                    /rpt_family="B1_MM"
repeat_region      18904..18931
                    /rpt_family="(TC)n"
repeat_region      18931..18931
                    /rpt_family="AT_rich"
repeat_region      18938..19077
                    /rpt_family="B1_MM"
repeat_region      19097..19165
                    /rpt_family="(MTE)"
repeat_region      19255..19462
                    /rpt_family="B3"
repeat_region      19981..20129
                    /rpt_family="MER70B"
repeat_region      20431..20580
                    /rpt_family="B1_MM"
repeat_region      20604..20727
                    /rpt_family="B1_MM"
repeat_region      20730..20933
                    /rpt_family="B2_Mm2"
repeat_region      21102..21150
                    /rpt_family="(TTTC)n"
repeat_region      21153..21301
                    /rpt_family="B1_MM"
repeat_region      21346..21553
                    /rpt_family="B2_Mm1"
repeat_region      21606..21803
                    /rpt_family="B2_Mm2"
repeat_region      21912..22026
                    /rpt_family="B1_MM"
repeat_region      22447..22628

```

```

Query Match      1.5%; Score 30; DB 10; Length 125105;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1715 GGCATCTCGATGGGCGCATCCGAGCATG 1744
         |||||
Db      2573 GGCATCTCGATGGGCGCATCCGAGCATG 2602

```

```

RESULT 44
AC108592/c      217688 bp      DNA      linear      HTG 15-NOV-2002
LOCUS      Rattus norvegicus clone CH230-144C19, *** SEQUENCING IN PROGRESS
DEFINITION
ACCESSION      AC108592
VERSION      AC108592.5 GI:25006698
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DNAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

```

```

REFERENCE
AUTHORS      1 (bases 1 to 217688)
Muzny,D,Marie, Metzker,M, Lee, A, Abramson, S, Adams, C, Alder, J,
Allen, C, Allen, H, Alsbrooks, S, Amin, A, Anguiano, D,
Anyalebechi, V, Aoyagi, A, Ayodeji, M, Baca, B, Baden, H,
Baldwin, D, Bandaranaike, D, Barber, M, Barnstead, M, Benahmed, F,
Bismal, K, Blair, J, Blankenburg, K, Blyth, P, Brown, M,
Bryant, N, Buhay, C, Burch, P, Burrell, K, Calderon, E,
Cardenas, V, Carter, K, Cavazos, I, Caesar, H, Center, A,
Chacko, J, Chavez, D, Chen, G, Chen, Y, Chen, Z, Chu, J,
Cleveland, C, Cockrell, R, Cox, C, Coyle, M, Cree, A, D'Souza, L,
Davila, M, L, Davis, C, Davy-Carroll, L, De Anda, C, Dederich, D,
Delgado, O, Denson, S, Deramo, C, Ding, Y, Dinh, H, Diya, K,
Draper, H, Dugan-Rocha, S, Dunn, A, Durbin, K, Duval, B, Eaves, K,
Egan, A, Escotto, M, Eugene, C, Evans, C, A, Falls, T, Fan, G,
Fernandez, S, Finley, M, Flagg, N, Forbes, L, Foster, M, Foster, P,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guervara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harrey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huiyk, S., Hume, J., Idibit, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mamoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okwunonu, G., Olafunso, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.D., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svetek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tindley, A., Tjof, S., Umani, K., Vals, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Woley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., Von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 217688)  
Worley, K.C.  
Direct Submission  
Submitted (31-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 217688)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23270241.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GPMS  
Center clone name: CH230-144C19  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 17651 bases at least Q40



Consensus quality: 182606 bases at least Q30  
 Consensus quality: 185802 bases at least Q20  
 Estimated insert size: 178707, sum-of-configs estimation  
 Quality coverage: 5x in Q20 bases; sum-of-configs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 configs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the configs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 161680: contig of 161680 bp in length  
 \* 161681 161780: gap of unknown length  
 \* 161781 170889: contig of 9109 bp in length  
 \* 170890 170989: gap of unknown length  
 \* 170990 177789: contig of 6800 bp in length  
 \* 177790 177889: gap of unknown length  
 \* 177890 189681: contig of 11792 bp in length  
 \* 189682 189781: gap of unknown length  
 \* 189782 211300: contig of 21519 bp in length  
 \* 211301 212789: contig of 1389 bp in length  
 \* 212790 212889: gap of unknown length  
 \* 212890 214330: contig of 1441 bp in length  
 \* 214331 214430: gap of unknown length  
 \* 214431 215739: contig of 1309 bp in length  
 \* 215740 215839: gap of unknown length  
 \* 215840 217688: contig of 1849 bp in length.

## FEATURES

source  
 1. 217688  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-144C19"  
 7490..10042  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 10093..11330  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 71480..74410  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 87246..89004  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 154433..155684  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 160679..161680  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 170990..172417  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 175878..177789  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 177890..179119  
 /note="wgs\_contig"  
 misc\_feature  
 /note="wgs\_contig"  
 189782..190939  
 /note="wgs\_contig"

## ORIGIN

Query Match 1.5%; Score 30; DB 2; Length 217688;  
 Best Local Similarity 100.0%; Pred. No. 0.00027;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 GGCACTCTGCACTGGCCATCCAGAGCATG 1744  
 |||||  
 DB 114306 GGCACTCTGCACTGGCCATCCAGAGCATG 114277  
 |||||  
 RESULT 45  
 AC120780/c AC120780 193047 bp DNA linear PRI 06-NOV-2003  
 DEFINITION Pan troglodytes clone rp43-22b16, complete sequence.

ACCESSION AC120780  
 VERSION AC120780.32 GI:38016068  
 KEYWORDS HTG.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.  
 1 (bases 1 to 193047)  
 Zhou, L., Fu, Y., Shi, R., Wu, J., Shaul, S., Eichler, E. and Roe, B.A.  
 Pan troglodytes BAC Clone rp43-22b16  
 Unpublished  
 2 (bases 1 to 193047)  
 Zhou, L., Fu, Y., Shi, R., Wu, J., Shaul, S., Eichler, E. and Roe, B.A.  
 Direct Submission  
 Submitted (09-MAY-2002) Department of Chemistry And Biochemistry,  
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
 3 (bases 1 to 193047)  
 Zhou, L., Fu, Y., Shi, R., Wu, J., Shaul, S., Eichler, E. and Roe, B.A.  
 Direct Submission  
 Submitted (29-OCT-2003) Department of Chemistry And Biochemistry,  
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
 4 (bases 1 to 193047)  
 Zhou, L., Fu, Y., Shi, R., Wu, J., Shaul, S., Eichler, E. and Roe, B.A.  
 Direct Submission  
 Submitted (06-NOV-2003) Department of Chemistry And Biochemistry,  
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
 On Oct 29, 2003 this sequence version replaced gi:37999296.

## COMMENT

----- Genome Center  
 Center: Department of Chemistry And Biochemistry  
 The University Of Oklahoma  
 Center code: UOKNOR

## FEATURES

source  
 1. 193047  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="rp43-22b16"  
 /clone\_lib="RPc1 - 43 Male Chimpanzee BAC Library"

## ORIGIN

Query Match 1.3%; Score 27; DB 9; Length 193047;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2010 ACCTACTAAATATATAAATTAGCTG 2036  
 |||||  
 DB 181367 ACCTACTAAATATATAAATTAGCTG 181341  
 |||||

RESULT 46  
 AC097332/c AC097332 194191 bp DNA linear HTG 15-JUN-2002  
 DEFINITION Pan troglodytes clone Rp43-22b16, WORKING DRAFT SEQUENCE, 5  
 unordered pieces.

ACCESSION AC097332  
 VERSION AC097332.3 GI:21397204  
 KEYWORDS HTG; HTGS; PHASE1; HTGS\_DRAFT; HTGS\_FULLP.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.  
 1 (bases 1 to 194191)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C.,  
 Albrooks, S.L., Amaralunge, H.C., Are, J.R., Ayale, M., Banks, T.,  
 Barbarella, J., Benson, J., Blumage, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douhwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Fratcz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homai, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.B., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C., Kravovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, Y., Louleaged, H., Lozdo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primas, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojudo, K., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shoshart, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umanil, K., Vaquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 194191)  
Worley, K.C.

Direct Submission  
Submitted (14-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 194191)  
Worley, K.C.

Direct Submission  
Submitted (15-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 13, 2002 this sequence version replaced gi:21217398.

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information  
Center project name: ZUN2  
Center clone name: RP43-22B16

Summary Statistics  
Sequencing vector: Plasmid, M77769  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Assembly program: Phrap, version 0.990329  
Consensus quality: 193448 bases at least Q40  
Consensus quality: 193860 bases at least Q30  
Consensus quality: 194094 bases at least Q20  
Estimated insert size: 194270; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-tp estimation  
Quality coverage: 8.5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 75617: contig of 75617 bp in length  
75618 75717: gap of unknown length  
75718 132426: contig of 56709 bp in length  
132427 132526: gap of unknown length  
132527 180942: contig of 48416 bp in length  
180943 181042: gap of unknown length  
181043 183092: contig of 2050 bp in length  
183093 183192: gap of unknown length  
183193 194191: contig of 10999 bp in length.

Location/Qualifiers  
1. 194191  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-22B16"

ORIGIN  
Query Match 1.3%; Score 27; DB 2; length 194191;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2010 ACCTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 65748 ACCTACTAAATATATAAATTAGCTG 65722  
|||||

RESULT 47  
AC148550  
LOCUS  
DEFINITION  
AC148550 206315 bp DNA linear HTG 07-APR-2004  
ordered pieces.  
AC148550  
VERSION  
AC148550.2 GI:46250780  
KEYWORDS  
HTG; HTGS PHASE2; HTGS DRAFT.  
SOURCE  
Callitrix jacchus (white-tufted-ear marmoset)  
ORGANISM  
Callitrix jacchus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;  
Callitrix.  
1 (bases 1 to 206315)  
Antonellis, A., Ayele, K., Benjamin, B., Blakesley, R.W.,  
Boutford, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,  
Coleman, H., Daki, N., Engle, J., Granice, S., Guan, X., Gupta, J.,  
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B.,  
Idol, J.R., Jones, C., Karlins, B., Kim, H., Kwong, P., Latic, P.,  
Larson, S., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,  
Margulies, E.H., Masiello, C., Maske, B., McDowell, J.,  
Mullikin, J.C., Paguirigan, C., Portnoy, M.E., Prasad, A., Puri, O.,  
Reddik-Dugue, N., Schandier, K., Schneider, M.G., Shah, K., Sison, C.,  
Stattin, P., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L.,  
Wetherly, K.D., Young, A. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 206315)  
Green, E.D.  
Direct Submission  
Submitted (18-MAR-2004) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 206315)  
Green, E.D.  
Direct Submission  
Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
On Apr 7, 2004 this sequence version replaced gi:45544611.

Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: [nisc\\_zoo@nhgri.nih.gov](mailto:nisc_zoo@nhgri.nih.gov)

## ----- Project Information

Center project name: fti  
Center clone name: 22821

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

## ----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 205372 bases at least Q40  
Consensus quality: 205683 bases at least Q30  
Consensus quality: 205766 bases at least Q20  
Insert size: 208000; agarose-fp  
Insert size: 205815; sum-of-contigs  
Quality coverage: 10.41x in Q20 bases; agarose-fp  
Quality coverage: 10.52x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and the accession number will be preserved.

1 9751: contig of 9751 bp in length  
\* 9752 9851: gap of unknown length  
\* 9852 25162: contig of 15311 bp in length  
\* 25163 25262: gap of unknown length  
\* 25263 27757: contig of 2495 bp in length  
\* 27758 27857: gap of unknown length  
\* 27858 43562: contig of 15705 bp in length  
\* 43563 43663: gap of unknown length  
\* 43663 147798: contig of 104136 bp in length  
\* 147799 147899: gap of unknown length  
\* 147899 206315: contig of 58417 bp in length.

## FEATURES

1.206315 Location/Qualifiers  
/organism="Callithrix jacchus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9483"  
/clone="CH259-222821"  
/clone\_1lb="CH259"  
/note="BAC resource: http://bacpac.chori.org/"  
1.9751  
/note="assembly\_fragment  
clone end:SP6  
vector side:left"  
misc\_feature 9852..25162  
/note="assembly\_fragment"  
misc\_feature 25263..27757  
/note="assembly\_fragment"  
misc\_feature 27858..43562  
/note="assembly\_fragment"  
misc\_feature 43663..147798  
/note="assembly\_fragment"  
misc\_feature 101350..206315  
/note="clone overlaps with GenBank Accession Number AC148552 clone CH259-314B19 (center project name fuz)"  
147899..206315  
/note="assembly\_fragment

## ORIGIN

clone end:T7  
vector\_side:right"

Query Match 1.3%; Score 27; DB 2; Length 206315;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2010 ACCTACTAAATATTAATTAATGCTG 2036  
141041 ACCTACTAAATATTAATTAATGCTG 141067

## RESULT 48

AC148552 212879 bp DNA 1linear HTG 07-APR-2004  
LOCUS Callithrix jacchus clone CH259-314B19, WORKING DRAFT SEQUENCE, 8

## DEFINITION

ordered pieces.

## ACCESSION

AC148552.2 GI:46250781

## KEYWORDS

HTG; HTGS\_PHASE2; HTGS\_DRAFT.

## SOURCE

Callithrix jacchus (white-tufted-ear marmoset)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## REFERENCE

1 (bases 1 to 212879)  
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;

## AUTHORS

Antonelli,A., Ayele,K., Benjamin,B., Blakeley,R.W.,  
Boutford,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,  
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,  
Hachiguchi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,  
Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Latic,P.,  
Larson,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,  
Marquies,E.H., Masello,C., Maskeri,B., McDowell,J.,  
Mulliken,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puti,O.,  
Reddik-Dugue,N., Schandler,K., Schaefer,M.G., Shah,K., Sison,C.,  
Stratipod,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,  
Wetherby,K.D., Young,A. and Green,E.D.  
NISC Comparative Sequencing Initiative

## TITLE

Unpublished

## JOURNAL

2 (bases 1 to 212879)

## REFERENCE

Submitted (18-MAR-2004) NIH Intramural Sequencing Center, 8717

## AUTHORS

Groveomont Circle, Gaithersburg, MD 20877, USA

## TITLE

3 (bases 1 to 212879)

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## AUTHORS

Groveomont Circle, Gaithersburg, MD 20877, USA

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## JOURNAL

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## REFERENCE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

## TITLE

Submitted (07-APR-2004) NIH Intramural Sequencing Center, 8717

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 210791 bases at least Q40  
 Consensus quality: 211551 bases at least Q30  
 Consensus quality: 211987 bases at least Q20  
 Insert size: 21400; agarose-fp  
 Insert size: 212179; sum-of-contigs  
 Quality coverage: 10.15x in Q20 bases; agarose-fp  
 Quality coverage: 10.23x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \*  
 \* 1 46270: contig of 46270 bp in length  
 \* 46271 46370: gap of unknown length  
 \* 46371 97514: contig of 51144 bp in length  
 \* 97515 100068: gap of unknown length  
 \* 100069 100168: contig of 2454 bp in length  
 \* 100169 123865: contig of 23697 bp in length  
 \* 123866 123965: gap of unknown length  
 \* 123966 185007: contig of 61042 bp in length  
 \* 185008 185107: gap of unknown length  
 \* 185108 188335: contig of 3228 bp in length  
 \* 188336 188435: gap of unknown length  
 \* 188436 194989: contig of 6554 bp in length  
 \* 194990 195089: gap of unknown length  
 \* 195090 212879: contig of 17790 bp in length.  
 Location/Qualifiers  
 1..212879  
 /organism="Callithrix jacchus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9483"  
 /clone\_lib="CH259-314B19"  
 /clone\_1ib="CH259"  
 /note="BAC resource: http://bacpac.choi.org/"  
 1..105030  
 /note="clone overlaps with GenBank Accession Number  
 AC148550 clone CH259-222B21 (center project name ftl.)"  
 1..46270  
 /note="assembly\_fragment  
 clone\_end:T7  
 vector\_side:left"  
 46371..97514  
 /note="assembly\_fragment"  
 97615..100068  
 /note="assembly\_fragment"  
 100169..123865  
 /note="assembly\_fragment"  
 123966..185007  
 /note="assembly\_fragment"  
 138656..212879  
 /note="clone overlaps with GenBank Accession Number  
 AC148627 clone CH259-214M19 (center project name fuy)"  
 185108..188335  
 /note="assembly\_fragment"  
 188436..194989  
 /note="assembly\_fragment"  
 195090..212879  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:right"

ORIGIN  
 Query Match 1.3%; Score 27; DB 2; Length 212879;  
 Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2010 ACCTACTAAAAATATAAATAGCTG 2036  
 Db 39509 ACCTACTAAAAATATAAATAGCTG 39535

RESULT 49  
 BV089370 577 bp DNA linear SRS 15-OCT-2003  
 LOCUS BV089370  
 DEFINITION RPMMSRQ0001266 Roche Palo Alto Mus musculus SRS genomic, sequence  
 tagged site.  
 ACCESSION BV089370  
 VERSION BV089370.1 GI:37666849  
 KEYWORDS SRS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 577)  
 Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,  
 McPherson,J.D., Roernzler,D. and Peltz,G.  
 Mus musculus SNPs  
 Unpublished (2003)  
 COMMENT  
 TITLE Contact: Jonathan Usuka  
 JOURNAL Roche Palo Alto Genetics and Genomics Department  
 3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA  
 Tel: 6508555807  
 Email: Jonathan.Usuka@roche.com  
 Primer A: No primer submitted  
 Primer B: No primer submitted.  
 Location/Qualifiers  
 1..577  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /map="10-480-9-CANA0111734.1.1.5628"  
 /clone\_lib="Roche Palo Alto"  
 /note="SNPs developed from assay sequences derived from 15  
 different strains-of mice (as of October 1, 2003). Those  
 strains include A/J, A/HeJ, -129/Sv, AKR/J, B10.D2-H2/oshJ,  
 BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, -CAST/Ei, DBA/2J,  
 MRL/MpJ, NZB/BinJ, NZW/LacJ, SPRET/Ei,..."  
 <1..>577

#### FEATURES

source  
 1..577  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /map="10-480-9-CANA0111734.1.1.5628"  
 /clone\_lib="Roche Palo Alto"  
 /note="SNPs developed from assay sequences derived from 15  
 different strains-of mice (as of October 1, 2003). Those  
 strains include A/J, A/HeJ, -129/Sv, AKR/J, B10.D2-H2/oshJ,  
 BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, -CAST/Ei, DBA/2J,  
 MRL/MpJ, NZB/BinJ, NZW/LacJ, SPRET/Ei,..."  
 <1..>577

#### ORIGIN

Query Match 1.3%; Score 26; DB 11; Length 577;  
 Best Local Similarity 100.0%; Pred. No. 0.093;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 GCGTCCCCCGAGGCTTCACGCCCT 321  
 Db 269 GCGTCCCCCGAGGCTTCACGCCCT 294

RESULT 50  
 CQ843504/c 3932 bp DNA linear PAT 02-AUG-2004  
 LOCUS CQ843504  
 DEFINITION Sequence 2151 from Patent EP1440981.  
 ACCESSION CQ843504  
 VERSION CQ843504.1 GI:50895291  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1  
 Isegai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,  
 Yamamoto,U., Isono,Y., Nagai,K. and Irie,R.  
 TITLE Full-length human cdna  
 JOURNAL Patent: EP 1440981-A 2151 28-JUL-2004;



COMMENT OK 73019, USA  
On Oct 30, 1998 this sequence version replaced gi:2967608.  
Because these overlapping clones came from different libraries.  
FEATURES  
source  
1.38715  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="22q11.2"  
/clone="98c4"  
/note="Left\_neighbor: 19d3; Right\_neighbor: 49c12"

ORIGIN

Query Match 1.3%; Score 26; DB 9; Length 38715;  
Best Local Similarity 100.0%; Pred No. 0.05;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2011 CCTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 21761 CCTACTAAATATATAAATTAGCTG 21736

RESULT 53  
AF293359 49087 bp DNA linear pri 14-OCT-2000  
LOCUS Homo sapiens desmocollin 3 (DSC3) gene, complete cds, alternatively  
DEFINITION spliced  
ACCESSION AF293359  
VERSION AF293359.1 GI:10801204  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 49087)  
AUTHORS Whitlock,N.V., Hunt,D.M., Rickman,L., Malhi,S., Vogazianou,A.P.,  
Dawson,L.F., Bady,R.A., Buxton,R.S. and McGrath,J.A.  
TITLE Genomic organization and amplification of the human desmosomal  
cadherin genes DSC1 and DSC3, encoding desmocollin types 1 and 3  
JOURNAL Biochem. Biophys. Res. Commun. 276 (2), 454-460 (2000)  
MEDLINE 20484176  
PUBMED 11027496  
REFERENCE 2 (bases 1 to 49087)  
AUTHORS Whitlock,N.V., Bady,R.A. and McGrath,J.A.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-2000) Cellular and Molecular Pathology, St John's  
Institute of Dermatology, St Thomas' Hospital, Lambeth Palace Road,  
London SE1 7EH, England  
FEATURES  
source  
1.49087  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18q12.1"  
1.49087  
/gene="DSC3"

gene  
mRNA  
mRNA  
exon  
1.81  
/product="desmocollin 3b"  
/gene="DSC3"  
/number=1

CDS

join(13..81,10397..10481,11500..11699,13041..13160,  
16758..16913,18179..18323,20170..20336,23872..24006,  
24416..24601,34163..34419,34530..34672,35556..35780,  
38322..38546,40946..41067,45631..45888,48179..48205)  
/gene="DSC3"  
/note="desmosomal cadherin"  
/codon\_start=1  
/product="desmocollin 3a"  
/protein\_id="AA023426.1"  
/db\_xref="GI:10801205"  
/translation="MAAGPRRSRGVAGVCHLLTLTVIFSRDGEACKKVILNVPEKLE  
ADKTIKIRVNLSECFRSADLIRNSDPDFVINDGSYYTARAVALSDKKSFTIWLQDKR  
KQTOKEGVVLLLEHQKVSXKTRHTRRETVLRARARWAPIPCSMQENSLDPPFLQDVE  
SDAAGNVTFYFISISRGVDKEPLNLFYIERDGNLCFTRPVDREYVFDJALYASTA  
DGSADNLPPLPIRVEDNDNHVPFTVAIVPEVLESRRPGFTVGVACVATDREBDPTM  
HTRLKYSILQOTPRSFGLSVHPSPGTVTTHSVLDRVDVKYSIMVVOIMDQCPFG  
LIGSTCIITITDSNDNAPTPRQNAVEAFVEBNANVEITLPIEDKULINTAMRVN  
FTILKGNENGHFKISTDKETNEGVLSVRPLNYEENRQVNLKIGVNNBAPFARDIPRV  
TALNRLATVAVRDLDEGECPAPQAVYRIKENLAVGSKINGYKAYADENRNGGLARY  
KLHDPKGMITIDEISGSIITSKILDREREYIPKELVNIYVLAIDKDRSCITGLAVN  
IEDVNDNPEILQEVVVICPKPMGYTDLAVDPDEPVHGAPEYFSLPMTSPISRLMS  
LTKVNDTARLSYQKNAQFORVTTIPITVKDASGAQATLRLVNLCECHPPOCRATRS  
STGVILGKRAIATILGIALFVSULLTIVCVFGRATKGRRPEDIQAQONLITSNEAP  
GDVRVCSANGFMTQTTNNSQGFPGTMSGMKNGQETITEMKGNQTLSECRGAGHH  
HTLDSCRGAGHTEVDCRYTSEHMSFTQPRLGEKILARNQENRMPSDQVYLYTYEG  
RSPSPAGSVGCCSEKQEBDLDFLNNLBKRFITLAEACTKR"  
join(13..81,10397..10481,11500..11699,13041..13160,  
16758..16913,18179..18323,20170..20336,23872..24006,  
24416..24601,34163..34419,34530..34672,35556..35780,  
38322..38546,40946..41067,45631..45888,48179..48205)  
/gene="DSC3"  
/note="desmosomal cadherin"  
/codon\_start=1  
/product="desmocollin 3b"  
/protein\_id="AA023427.1"  
/db\_xref="GI:10801206"  
/translation="MAAGPRRSRGVAGVCHLLTLTVIFSRDGEACKKVILNVPEKLE  
ADKTIKIRVNLSECFRSADLIRNSDPDFVINDGSYYTARAVALSDKKSFTIWLQDKR  
KQTOKEGVVLLLEHQKVSXKTRHTRRETVLRARARWAPIPCSMQENSLDPPFLQDVE  
SDAAGNVTFYFISISRGVDKEPLNLFYIERDGNLCFTRPVDREYVFDJALYASTA  
DGSADNLPPLPIRVEDNDNHVPFTVAIVPEVLESRRPGFTVGVACVATDREBDPTM  
HTRLKYSILQOTPRSFGLSVHPSTGVTITSHYLDREVDKYSILMKVQDDQCPFG  
LIGSTCIITVDSNDNAPTPRQNAVEAFVEBNANVEILRIPIEDKULINTAMRVN  
FTILKGNENGHFKISTDKETNEGVLSVRPLNYEENRQVNLKIGVNNBAPFARDIPRV  
TALNRLATVAVRDLDEGECPAPQAVYRIKENLAVGSKINGYKAYADENRNGGLARY  
KLHDPKGMITIDEISGSIITSKILDREREYIPKELVNIYVLAIDKDRSCITGLAVN  
IEDVNDNPEILQEVVVICPKPMGYTDLAVDPDEPVHGAPEYFSLPMTSPISRLMS  
LTKVNDTARLSYQKNAQFORVTTIPITVKDRAQQAATLRLVNLCECHPPOCRATRS  
STGVILGKRAIATILGIALFVSULLTIVCGVFGATKGRRPEDIQAQONLITSNEAP  
GDVRVCSANGFMTQTTNNSQGFPGTMSGMKNGQETITEMKGNQTLSECRGAGHH  
HTLDSCRGAGHTEVDCRYTSEHMSFTQPRLGEESIRHTTG"  
10397..10481  
/gene="DSC3"  
/number=2  
11500..11699  
/gene="DSC3"  
/number=3  
13041..13160  
/gene="DSC3"  
/number=4  
16758..16913  
/gene="DSC3"  
/number=5  
18179..18323  
/gene="DSC3"  
/number=6  
20170..20336  
/gene="DSC3"  
/number=7  
23872..24006  
/gene="DSC3"  
/number=8  
24416..24601

exon  
exon  
exon  
exon  
exon

```

/ gene="DSC3"
/ number=9
exon 34163..34419
/ gene="DSC3"
/ number=10
exon 34530..34672
/ gene="DSC3"
/ number=11
exon 35556..35780
/ gene="DSC3"
/ number=12
exon 38322..38546
/ gene="DSC3"
/ number=13
exon 40946..41067
/ gene="DSC3"
/ number=14
exon 45631..45888
/ gene="DSC3"
/ number=15
exon 48179..48221
/ gene="DSC3"
/ number=16
exon 48303..49087
/ gene="DSC3"
/ number=17

ORIGIN
Query Match 1.3%; Score 26; DB 9; Length 49087;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATAGCTG 2036
Db 9892 CCTACTAAATATATAAATAGCTG 9917
|||||
|||||

RESULT 54
AC132197/c
LOCUS AC132197
DEFINITION Homo sapiens chromosome 11 clone RP11-452112 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC132197
AC132197.1 GI:22549815
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 65608)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-452112
Unpublished
2 (bases 1 to 65608)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barina,W., Baerlein,V., Bloom,T., Boguski,M., Bouhassira,D.,
Cammarata,V., Chang,J., Chazaro,B., Choquet,Y., Collymore,A.,
Cook,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Garavito,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horion,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karcas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Mathews,C.,
McCarthy,M., Meldrum,J., Meneue,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retra,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

```

## JOURNAL COMMENT

Submitted (30-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L28114  
 Center clone name: 452\_I\_12

-----

\* NOTE: This record contains 81 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

```

1
688: contig of 688 bp in length
689
789: gap of 100 bp
789
1492: contig of 704 bp in length
1493
1593: gap of 100 bp
1593
2304: contig of 712 bp in length
2305
2404: gap of 100 bp
2405
3111: contig of 707 bp in length
3112
3211: gap of 100 bp
3212
3928: contig of 717 bp in length
3929
4028: gap of 100 bp
4029
4758: contig of 730 bp in length
4759
4858: gap of 100 bp
4859
5578: contig of 720 bp in length
5579
5678: gap of 100 bp
5679
6395: contig of 717 bp in length
6396
6495: gap of 100 bp
6496
7212: contig of 717 bp in length
7213
7312: gap of 100 bp
7313
8028: contig of 716 bp in length
8029
8128: gap of 100 bp
8129
8857: contig of 729 bp in length
8858
8957: gap of 100 bp
8958
9662: contig of 705 bp in length
9663
9762: gap of 100 bp
9763
10469: contig of 707 bp in length
10470
10569: gap of 100 bp
10570
11275: contig of 706 bp in length
11276
11375: gap of 100 bp
11376
12101: contig of 726 bp in length
12102
12201: gap of 100 bp
12202
12818: contig of 717 bp in length
12819
13018: gap of 100 bp
13019
13710: contig of 692 bp in length
13711
13810: gap of 100 bp
13811
14524: contig of 714 bp in length
14525
14624: gap of 100 bp
14625
15348: contig of 724 bp in length
15349
15448: gap of 100 bp
15449
16130: contig of 682 bp in length
16131
16230: gap of 100 bp
16231
16904: contig of 674 bp in length
16905
17004: gap of 100 bp
17005
17701: contig of 697 bp in length
17702
17801: gap of 100 bp
17802
18518: contig of 717 bp in length
18519
18618: gap of 100 bp
18619
19331: contig of 713 bp in length
19332
19431: gap of 100 bp
19432
19431: gap of 100 bp
19431

```

\* 19432 20131: contig of 700 bp in length  
\* 20132 20231: gap of 100 bp  
\* 20232 20933: contig of 702 bp in length  
\* 20934 21033: gap of 100 bp  
\* 21034 21743: contig of 710 bp in length  
\* 21744 21843: gap of 100 bp  
\* 21844 22563: contig of 720 bp in length  
\* 22564 22663: gap of 100 bp  
\* 22664 23381: contig of 718 bp in length  
\* 23382 23481: gap of 100 bp  
\* 23482 24183: contig of 702 bp in length  
\* 24184 24283: gap of 100 bp  
\* 24284 24979: contig of 696 bp in length  
\* 24980 25079: gap of 100 bp  
\* 25080 25806: contig of 727 bp in length  
\* 25807 25906: gap of 100 bp  
\* 25907 26601: contig of 695 bp in length  
\* 26602 26701: gap of 100 bp  
\* 26702 27416: contig of 715 bp in length  
\* 27417 27516: gap of 100 bp  
\* 27517 28224: contig of 708 bp in length  
\* 28225 28325: gap of 100 bp  
\* 28325 29023: contig of 699 bp in length  
\* 29024 29123: gap of 100 bp  
\* 29124 29839: contig of 716 bp in length  
\* 29840 29939: gap of 100 bp  
\* 29940 30650: contig of 711 bp in length  
\* 30651 30750: gap of 100 bp  
\* 30751 31462: contig of 712 bp in length  
\* 31463 31562: gap of 100 bp  
\* 31563 32274: contig of 712 bp in length  
\* 32275 32374: gap of 100 bp  
\* 32375 33093: contig of 719 bp in length  
\* 33094 33193: gap of 100 bp  
\* 33194 33910: contig of 717 bp in length  
\* 33911 34010: gap of 100 bp  
\* 34011 34721: contig of 711 bp in length  
\* 34722 34821: gap of 100 bp  
\* 34822 35545: contig of 724 bp in length  
\* 35546 35645: gap of 100 bp  
\* 35646 36356: contig of 711 bp in length  
\* 36357 36456: gap of 100 bp  
\* 36457 37175: contig of 719 bp in length  
\* 37176 37275: gap of 100 bp  
\* 37276 38005: contig of 730 bp in length  
\* 38006 38105: gap of 100 bp  
\* 38106 38822: contig of 717 bp in length  
\* 38823 38922: gap of 100 bp  
\* 38923 39642: contig of 720 bp in length  
\* 39643 39742: gap of 100 bp  
\* 39743 40463: contig of 721 bp in length  
\* 40464 40563: gap of 100 bp  
\* 40564 41293: contig of 730 bp in length  
\* 41294 41393: gap of 100 bp  
\* 41394 42102: contig of 709 bp in length  
\* 42103 42202: gap of 100 bp  
\* 42203 42887: contig of 685 bp in length  
\* 42888 42987: gap of 100 bp  
\* 42988 43705: contig of 718 bp in length  
\* 43706 43805: gap of 100 bp  
\* 43806 44516: contig of 711 bp in length  
\* 44517 44616: gap of 100 bp  
\* 44617 45327: contig of 711 bp in length  
\* 45328 45427: gap of 100 bp  
\* 45428 46150: contig of 723 bp in length  
\* 46151 46250: gap of 100 bp  
\* 46251 46973: contig of 723 bp in length  
\* 46974 47073: gap of 100 bp  
\* 47074 47782: contig of 709 bp in length  
\* 47783 47882: gap of 100 bp  
\* 47883 48595: contig of 713 bp in length  
\* 48596 48695: gap of 100 bp  
\* 48696 49391: contig of 696 bp in length

Query Match 1.3% Score 26; DB 2; Length 65608;  
Best Local Similarity 100.0%; Fred. No. 0.046;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 2011 CCTACTAAATATATAATAGCTG 2036  
Db 47659 CCTACTAAATATATAATAGCTG 47634  
|||||  
RESULT 55  
AC073517 67706 bp DNA linear PRI 08-OCT-2003  
LOCUS Homo sapiens BAC clone RP11-163B9 from 7, complete sequence.  
AC073517  
AC073517 AC073517.6 GI:18482312  
VERSION  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (baaes 1 to 67706)  
Suleston,J.E. and Wilson,R.  
Toward a complete human genome sequence  
Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074  
REFERENCE  
2 (baaes 1 to 67706)  
Haakenson,W., Hawkins,N., Mclellam,M. and Nicol,C.  
The sequence of Homo sapiens BAC clone RP11-163B9  
Unpublished (2001)  
JOURNAL  
3 (baaes 1 to 67706)  
Waterston,R.H.  
Direct Submission  
Submitted (21-JUN-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
4 (baaes 1 to 67706)  
Waterston,R.H.  
Direct Submission  
Submitted (06-JUN-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
5 (baaes 1 to 67706)  
Waterston,R.H.  
Direct Submission  
Submitted (03-JUL-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
6 (baaes 1 to 67706)  
Waterston,R.  
Direct Submission  
Submitted (07-NOV-2001) Department of Genetics, Washington



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
7 (bases 1 to 67706)  
Waterston,R.H.  
Direct Submission  
Submitted (03-FEB-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
8 (bases 1 to 67706)  
Waterston,R.  
Direct Submission  
Submitted (21-FEB-2002) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
9 (bases 1 to 67706)  
Waterston,R.  
Direct Submission  
Submitted (29-APR-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
10 (bases 1 to 67706)  
Wilson,R.  
Direct Submission  
Submitted (08-OCT-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Feb 3, 2002 this sequence version replaced g1:14318396.  
----- Genome Center  
-----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: saplens@wustl.wustl.edu  
----- Summary Statistics  
-----  
Center project name: H\_NH0163E09  
-----  
NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.  
  
This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.  
  
MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and  
sequencing collaboration between the NHGRI Chromosome 7 Mapping  
Project (Eric D. Green, Director), John D. McPherson in the  
Department of Genetics (Washington University), and the Washington  
University Genome Sequencing Center. For additional information  
about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send  
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu  
  
SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male  
donor, as described by Osogawa,K., Moon,P.Y., Zhao,B., Frengen,B.,  
Tateno,M., Catalanese,J.J. and de Jong,P.J. (1998) An improved  
approach for construction of bacterial artificial chromosome  
libraries. Genomics 51:1-8. The clone may be obtained either from  
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong  
and coworkers at http://www.chori.org  
VECTOR: pBACe3.6  
  
NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the right is AF030453, 200 bp overlap.  
Actual start of this clone is at base position 1 of RP11-163B9  
actual end is at base position 53581 of CTA-313A17.  
Location/Qualifiers  
1..67706  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7"  
/clone="RP11-163B9"  
/clone\_11b="RPCI-11"  
594..858  
/rpt\_family="Alu"  
859..883  
repeat\_region  
/rpt\_family="(CA)n"  
884..897  
repeat\_region  
/rpt\_family="Alu"  
1190..1375  
repeat\_region  
/rpt\_family="(TTCC)n"  
1376..1690  
repeat\_region  
/rpt\_family="Alu"  
1703..1935  
repeat\_region  
/rpt\_family="Alu"  
2711..2842  
repeat\_region  
/rpt\_family="L2"  
2844..3154  
repeat\_region  
/rpt\_family="Alu"  
3155..3439  
repeat\_region  
/rpt\_family="Alu"  
3769..4061  
repeat\_region  
/rpt\_family="Alu"  
4062..4213  
repeat\_region  
/rpt\_family="Alu"  
4349..4640  
repeat\_region  
/rpt\_family="Alu"  
5069..5176  
repeat\_region  
/rpt\_family="MIR"  
5755..5877  
repeat\_region  
/rpt\_family="Alu"  
8040..8284  
repeat\_region  
/rpt\_family="Alu"  
8322..8560  
repeat\_region  
/rpt\_family="MIR"  
8589..8880  
repeat\_region  
/rpt\_family="Alu"  
9023..9224  
repeat\_region  
/rpt\_family="MIR2\_type"  
9225..9359  
repeat\_region  
/rpt\_family="Alu"  
9360..9660  
repeat\_region  
/rpt\_family="Alu"  
9661..9857  
repeat\_region  
/rpt\_family="Alu"  
9858..10190  
repeat\_region  
/rpt\_family="MIR2\_type"  
10191..10484  
repeat\_region  
/rpt\_family="Alu"  
10485..10554  
repeat\_region  
/rpt\_family="Alu"  
10556..10739  
repeat\_region  
/rpt\_family="Alu"  
10743..11297  
repeat\_region  
/rpt\_family="MIR2\_type"  
11553..11815  
repeat\_region  
/rpt\_family="L1"  
11816..12104  
repeat\_region  
/rpt\_family="Alu"  
12105..12163  
repeat\_region  
/rpt\_family="(CA)n"  
12170..12578  
repeat\_region  
/rpt\_family="L1"  
12579..12869  
repeat\_region  
/rpt\_family="Alu"  
12870..13006  
repeat\_region  
/rpt\_family="L1"  
13007..13124  
repeat\_region  
/rpt\_family="Alu"  
13125..13651  
repeat\_region  
/rpt\_family="L1"

```
repeat_region 13652..13966
                /rpt_family="Alu"
repeat_region 13967..14079
                /rpt_family="L1"
repeat_region 14080..14380
                /rpt_family="Alu"
repeat_region 14381..14791
                /rpt_family="L1"
repeat_region 14795..14846
                /rpt_family="Alu"
repeat_region 14847..15146
                /rpt_family="Alu"
repeat_region 15147..15317
```

```
Query Match      1.3%; Score 26; DB 9; Length 67706;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2011 CCTACTAAATATATAAATAGCTG 2036
Db      20431 CCTACTAAATATATAAATAGCTG 20456
```

```
RESULT 56
AC135170      AC135170      69586 bp      DNA      linear      HTG 08-OCT-2002
LOCUS      Homo sapiens clone CTD-2242E9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION      AC135170
ACCESSION      AC135170
VERSION      AC135170.1 GI:23592129
KEYWORDS      HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
```

```
REFERENCE
AUTHORS      Birren,B., Nussbaum,C. and Lander,E.
TITLE      Homo sapiens, clone CTD-2242E9
JOURNAL      Unpublished
REFERENCE
AUTHORS      2 (bases 1 to 69586)
            2 (bases 1 to 69586)
```

Bairn,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,  
Cammarata,U., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horion,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karstae,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Menene,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retra,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zemek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (08-OCT-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L28327

Center clone name: 2242\_E\_9

-----

\* NOTE: This record contains 86 individual

\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 703: contig of 703 bp in length  
704 803: gap of 100 bp  
804 1510: contig of 707 bp in length  
1511 1610: gap of 100 bp  
1611 2316: contig of 706 bp in length  
2317 2416: gap of 100 bp  
2417 3120: contig of 704 bp in length  
3121 3220: gap of 100 bp  
3221 3928: contig of 708 bp in length  
3929 4028: gap of 100 bp  
4029 4744: contig of 716 bp in length  
4745 4844: gap of 100 bp  
4845 5551: contig of 707 bp in length  
5552 5651: gap of 100 bp  
5652 6370: contig of 719 bp in length  
6371 6470: gap of 100 bp  
6471 7188: contig of 718 bp in length  
7189 7288: gap of 100 bp  
7289 7992: contig of 704 bp in length  
7993 8092: gap of 100 bp  
8093 8795: contig of 703 bp in length  
8796 8895: gap of 100 bp  
8896 9600: contig of 705 bp in length  
9601 9700: gap of 100 bp  
9701 10422: contig of 722 bp in length  
10423 10522: gap of 100 bp  
10523 11227: contig of 705 bp in length  
11228 11327: gap of 100 bp  
11328 12030: contig of 703 bp in length  
12031 12130: gap of 100 bp  
12131 12835: contig of 705 bp in length  
12836 12935: gap of 100 bp  
12936 13657: contig of 722 bp in length  
13658 13757: gap of 100 bp  
13758 14487: contig of 730 bp in length  
14488 14587: gap of 100 bp  
14588 15311: contig of 724 bp in length  
15312 15411: gap of 100 bp  
15412 16126: contig of 715 bp in length  
16127 16226: gap of 100 bp  
16227 16957: contig of 731 bp in length  
16958 17057: gap of 100 bp  
17058 17754: contig of 697 bp in length  
17755 17854: gap of 100 bp  
17855 18570: contig of 716 bp in length  
18571 18670: gap of 100 bp  
18671 19366: contig of 696 bp in length  
19367 19466: gap of 100 bp  
19467 20186: contig of 720 bp in length  
20187 20286: gap of 100 bp  
20287 20999: contig of 713 bp in length  
21000 21099: gap of 100 bp  
21100 21803: contig of 704 bp in length  
21804 21903: gap of 100 bp  
21904 22625: contig of 722 bp in length  
22626 22725: gap of 100 bp  
22726 23446: contig of 721 bp in length  
23447 23546: gap of 100 bp  
23547 24249: contig of 703 bp in length  
24249 24349: gap of 100 bp  
24350 25064: contig of 715 bp in length  
25065 25164: gap of 100 bp  
25165 25874: contig of 710 bp in length

```

* 25974: gap of 100 bp
* 25975: contig of 707 bp in length
* 26681: gap of 100 bp
* 26682: contig of 705 bp in length
* 27486: gap of 100 bp
* 27487: contig of 710 bp in length
* 28296: gap of 100 bp
* 28297: contig of 701 bp in length
* 29098: gap of 100 bp
* 29197: contig of 725 bp in length
* 29922: gap of 100 bp
* 29923: contig of 722 bp in length
* 30023: gap of 100 bp
* 30744: contig of 722 bp in length
* 30844: gap of 100 bp
* 31566: contig of 722 bp in length
* 31567: gap of 100 bp
* 31667: contig of 712 bp in length
* 32378: gap of 100 bp
* 32479: contig of 703 bp in length
* 33181: gap of 100 bp
* 33182: contig of 706 bp in length
* 33987: gap of 100 bp
* 34088: contig of 695 bp in length
* 34782: gap of 100 bp
* 34883: contig of 720 bp in length
* 35603: gap of 100 bp
* 35702: contig of 710 bp in length
* 36412: gap of 100 bp
* 36512: contig of 708 bp in length
* 37220: gap of 100 bp
* 37221: contig of 703 bp in length
* 38023: gap of 100 bp
* 38123: contig of 722 bp in length
* 38845: gap of 100 bp
* 38945: contig of 732 bp in length
* 39677: gap of 100 bp
* 39778: contig of 725 bp in length
* 40502: gap of 100 bp
* 41323: contig of 721 bp in length
* 41423: gap of 100 bp
* 42140: contig of 717 bp in length
* 42240: gap of 100 bp
* 42952: contig of 712 bp in length
* 43052: gap of 100 bp
* 43053: contig of 702 bp in length
* 43755: gap of 100 bp
* 43854: contig of 706 bp in length
* 44560: gap of 100 bp
* 44561: contig of 723 bp in length
* 45383: gap of 100 bp
* 45483: contig of 716 bp in length
* 46199: gap of 100 bp
* 46299: contig of 693 bp in length
* 46992: gap of 100 bp
* 47092: contig of 698 bp in length
* 47790: gap of 100 bp
* 47791: contig of 718 bp in length
* 48608: gap of 100 bp
* 48708: contig of 722 bp in length
* 49430: gap of 100 bp
* 49431: contig of 728 bp in length
* 50258: gap of 100 bp
* 50358: contig of 724 bp in length
* 51082: gap of 100 bp
* 51182: contig of 706 bp in length
* 51888: gap of 100 bp
* 51988: contig of 705 bp in length
* 52693: gap of 100 bp
* 52793: contig of 712 bp in length
* 52994: gap of 100 bp
* 53505: gap of 100 bp
* 53605: contig of 716 bp in length
* 54321: gap of 100 bp
* 54421: contig of 707 bp in length
* 55128: gap of 100 bp
* 55129: contig of 707 bp in length

```

```

* 55229 55913: contig of 685 bp in length
* 55914 56013: gap of 100 bp
* 56014 56728: contig of 715 bp in length
* 56729 56828: gap of 100 bp
* 56829 57550: contig of 722 bp in length
* 57551 57650: gap of 100 bp

Query Match 1.3%; Score 26; DB 2; Length 69586;
Best Local Similarity 100.0%; Fred. No. 0.046;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAATAGCTG 2036
DB 38366 CCTACTAAATATATAATAGCTG 38391
|||||

```

```

RESULT 57
AL391378/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-173D22 on chromosome 10,
complete sequence.
ACCESSION AL391378
VERSION AL391378.10 GI:14596356
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 73421)
AUTHORS Heath, P.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

```

## COMMENT

On Jul 4, 2001 this sequence version replaced GI:11190595. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SwissProt, Tr: TrEMBL, Wp: WormPeP, Information on the WormPeP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10> RP11-173D22 is from the library RDCI-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

## FEATURES

```

source
1. 73421
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="10"
   /clone="RP11-173D22"

```

repeat\_region /clone\_11b="Rpci-11.1"  
42. .341  
/note="Aluub repeat: matches 1. .298 of consensus"  
repeat\_region 375. .681  
/note="AluY repeat: matches 1. .307 of consensus"  
repeat\_region 933. .1244  
/note="AluX repeat: matches 1. .311 of consensus"  
repeat\_region 1421. .1721  
/note="MER33 repeat: matches 2. .333 of consensus"  
repeat\_region 2476. .2521  
/note="LTR43 repeat: matches 211. .257 of consensus"  
repeat\_region 2530. .3104  
/note="L1P48 repeat: matches 5588. .6162 of consensus"  
repeat\_region 3107. .3414  
/note="AluX repeat: matches 1. .301 of consensus"  
repeat\_region 3582. .3643  
/note="31 copies 2 mer tg 80% conserved"  
repeat\_region 4308. .4612  
/note="AluX repeat: matches 1. .305 of consensus"  
repeat\_region 4674. .4971  
/note="AluX repeat: matches 1. .299 of consensus"  
repeat\_region 6798. .6856  
/note="MIR repeat: matches 48. .105 of consensus"  
repeat\_region 7068. .7393  
/note="AluU repeat: matches 1. .302 of consensus"  
repeat\_region 7526. .7954  
/note="MSTC repeat: matches 1. .405 of consensus"  
repeat\_region 8561. .8924  
/note="MLT1A1 repeat: matches 1. .365 of consensus"  
repeat\_region 9657. .9680  
/note="12 copies 2 mer tt 100% conserved"  
repeat\_region 9686. .9863  
/note="MLT1A2 repeat: matches 210. .388 of consensus"  
repeat\_region 9870. .10457  
/note="MER61B repeat: matches 7. .565 of consensus"  
repeat\_region 10459. .10684  
/note="MLT1A2 repeat: matches 1. .198 of consensus"  
repeat\_region 11377. .11495  
/note="MIR repeat: matches 30. .145 of consensus"  
repeat\_region 13000. .13115  
/note="MERSA repeat: matches 72. .189 of consensus"  
repeat\_region 13590. .13825  
/note="MER30 repeat: matches 1. .230 of consensus"  
repeat\_region 14327. .14666  
/note="MER1B repeat: matches 1. .337 of consensus"  
repeat\_region 14717. .14944  
/note="MER7A repeat: matches 117. .345 of consensus"  
repeat\_region 14945. .15027  
/note="MADE1 repeat: matches 1. .80 of consensus"  
repeat\_region 15028. .15147  
/note="MER7A repeat: matches 1. .117 of consensus"  
repeat\_region 15940. .16237  
/note="L1ME repeat: matches 5470. .5759 of consensus"  
repeat\_region 17584. .17714  
/note="MIR repeat: matches 50. .192 of consensus"  
repeat\_region 17837. .17927  
/note="MIR repeat: matches 157. .256 of consensus"  
repeat\_region 19430. .19719  
/note="AluSq repeat: matches 6. .297 of consensus"  
repeat\_region 19824. .19942  
/note="MIR repeat: matches 52. .175 of consensus"  
repeat\_region 20384. .20694  
/note="AluX repeat: matches 1. .311 of consensus"  
repeat\_region 21292. .21874  
/note="L1MC1 repeat: matches 5733. .6327 of consensus"  
repeat\_region 26049. .26093  
/note="MER47 repeat: matches 2251. .2295 of consensus"  
repeat\_region 26093. .26171  
/note="MER47 repeat: matches 2240. .2333 of consensus"  
repeat\_region 26207. .26389  
/note="MIR repeat: matches 60. .242 of consensus"  
repeat\_region 26390. .26764  
/note="THE1B repeat: matches 1. .364 of consensus"

repeat\_region 26765. .26823  
/note="MIR repeat: matches 2. .60 of consensus"  
repeat\_region 27124. .27422  
/note="AluX repeat: matches 1. .296 of consensus"  
repeat\_region 27427. .27602  
/note="MERSA repeat: matches 1. .189 of consensus"  
repeat\_region 33086. .33180  
/note="TIGER1 repeat: matches 72. .162 of consensus"  
repeat\_region 33510. .33597  
/note="L12 repeat: matches 2606. .2708 of consensus"  
repeat\_region 33665. .33963  
/note="AluSq repeat: matches 1. .299 of consensus"  
repeat\_region 34218. .34380  
/note="MIR repeat: matches 58. .232 of consensus"  
repeat\_region 35080. .35186  
/note="L12 repeat: matches 2592. .2703 of consensus"  
repeat\_region 35454. .35521  
/note="2 copies 34 mer 100% conserved"  
repeat\_region 36770. .37146  
/note="L1M1 repeat: matches 989. .1366 of consensus"  
repeat\_region 37213. .37590  
/note="L1M1 repeat: matches 1369. .1450 of consensus"  
repeat\_region 37609. .37904  
/note="AluX repeat: matches 1. .294 of consensus"  
repeat\_region 37906. .38374  
/note="L1 repeat: matches 2348. .2821 of consensus"  
repeat\_region 38398. .38612  
/note="L1MA8 repeat: matches 6067. .6288 of consensus"  
repeat\_region 38648. .38792  
/note="MIR repeat: matches 10. .173 of consensus"  
repeat\_region 40066. .40375  
/note="MER33 repeat: matches 1. .324 of consensus"  
repeat\_region 40456. .40690  
/note="MLT2D repeat: matches 1. .238 of consensus"  
repeat\_region 40680. .40780  
/note="MLT2D repeat: matches 453. .546 of consensus"  
repeat\_region 41211. .41380  
/note="5 copies 34 mer 68% conserved"  
repeat\_region 41509. .42296  
/note="L1MB5 repeat: matches 5354. .6137 of consensus"  
repeat\_region 42927. .42998  
/note="36 copies 2 mer tc 72% conserved"  
repeat\_region 43646. .44086  
/note="Charlie4 repeat: matches 21. .508 of consensus"  
repeat\_region 44146. .44386  
/note="MER47 repeat: matches 2083. .2323 of consensus"  
repeat\_region 44388. .44477  
/note="MER47 repeat: matches 1. .90 of consensus"  
repeat\_region 44478. .44804  
/note="MLT1A2 repeat: matches 2. .342 of consensus"  
repeat\_region 44819. .44978  
/note="L12 repeat: matches 2452. .2639 of consensus"  
repeat\_region 44979. .45302  
/note="MERS8B repeat: matches 1. .341 of consensus"  
repeat\_region 45303. .45413  
/note="L12 repeat: matches 2639. .2746 of consensus"  
repeat\_region 48649. .48752  
/note="2 copies 52 mer 96% conserved"  
repeat\_region 49853. .49894  
/note="21 copies 2 mer at 76% conserved"  
repeat\_region 51034. .51200  
/note="MERS45 repeat: matches 1. .171 of consensus"  
repeat\_region 51202. .51812  
/note="MER39B repeat: matches 13. .576 of consensus"  
repeat\_region 52156. .52461  
/note="Aluub repeat: matches 1. .300 of consensus"  
repeat\_region 55677. .55768  
/note="MERSA repeat: matches 2. .95 of consensus"  
repeat\_region 55842. .55888  
/note="MIR repeat: matches 41. .82 of consensus"  
repeat\_region 55889. .57194  
/note="HSMK1 repeat: matches 1. .1286 of consensus"

Query Match 1.3%; Score 26; DB 9; Length 73421;  
 Best Local Similarity 100.0%; Pred. No. 0.046;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
 |||||  
 DB 4498 CCTACTAAATATATAAATTAGCTG 4473

RESULT 58  
 HS452H17 99682 bp DNA linear PRI 05-MAR-2003  
 LOCUS Human DNA sequence from clone RP3-452H17 on chromosome Xq22.1-23,  
 DEFINITION complete sequence.  
 ACCESSION Z96810 GI:2276315  
 VERSION Z96810.1 GI:2276315  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 99682)  
 AUTHORS Grafham,D.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
 On Jul 24, 1997 this sequence version replaced gi:2198470.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 -----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/ChX RP3-452H17 is from the library RPc1-3 constructed by the group of Pletzer de Jong. For further details see http://www.chori.org/bacpac/home.htm  
 VECTOR: pCYPAC2.

FEATURES  
 source location/Qualifiers

1..99682  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="RZPD:RPc1P704H17452"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /map="q22.1-23"  
 /clone="RP3-452H17"  
 /clone\_11b="RPc1-3"

ORIGIN

Query Match 1.3%; Score 26; DB 9; Length 99682;  
 Best Local Similarity 100.0%; Pred. No. 0.044;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
 |||||  
 DB 24401 CCTACTAAATATATAAATTAGCTG 24426

RESULT 59  
 AC105430 105731 bp DNA linear PRI 05-APR-2003  
 LOCUS Homo sapiens chromosome 16 clone RP11-2C15, complete sequence.  
 DEFINITION AC105430  
 ACCESSION AC105430.2 GI:29568029  
 VERSION AC105430.2 GI:29568029  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 105731)  
 AUTHORS Alamos National Laboratory.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 105731)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-JAN-2002) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 REFERENCE 3 (bases 1 to 105731)  
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
 Alamos National Laboratory.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-APR-2003) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Apr 5, 2003 this sequence version replaced gi:18071325.  
 Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center and Los Alamos  
 National Laboratory  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 99.6% of Sequence;  
 Estimated Total Number of Errors is 0.3.  
 NOTE: This is not the entire sequence of the clone. It is clipped  
 over the overlaps with AC009122 and AC092379. The number of bases  
 overlapped with AC009122 is 4630 and with AC092379 is 20988.

FEATURES  
 source location/Qualifiers

1..105731  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 /clone="RP11-2C15"

ORIGIN

Query Match 1.3%; Score 26; DB 9; Length 105731;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
 |||||  
 DB 25224 CCTACTAAATATATAAATTAGCTG 25199

RESULT 60  
 AP000681/c 109149 bp DNA linear HTG 30-MAY-2000  
 LOCUS Homo sapiens chromosome 11 clone CMB9-2L13 map 11q23, WORKING DRAFT  
 DEFINITION SEQUENCE, 13 unordered pieces.  
 ACCESSION AP000681  
 VERSION AP000681.3 GI:8118869  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 109149)  
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 TITLE Homo sapiens 109,149 genomic DNA of 11q23  
 JOURNAL Published Only in Database (1999)  
 REFERENCE 2 (bases 1 to 109149)  
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-NOV-1999) Maashita Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,  
 Japan (E-mail:hattori@gsc.riken.go.jp,  
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,  
 Fax:81-42-778-9924)  
 On May 31, 2000 this sequence version replaced gi:697555.  
 COMMENT ----- Genome Center  
 Center: RIKEN Genomic Sciences Center(GSC)  
 Center code: RIKEN  
 Web site: http://hgp.gsc.riken.go.jp/  
 Contact: hattori@gsc.riken.go.jp  
 ----- Project Information  
 Project name: HumDraFtl1  
 Center clone name: CMB9-2L13  
 ----- Summary Statistics  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 99833 bases at least Q40  
 Consensus quality: 104437 bases at least Q30  
 Consensus quality: 106636 bases at least Q20  
 Insert size: 107949; sum-of-contents  
 Quality coverage: 4.90x in Q20 bases; sum-of-contents  
 -----

NOTE: This is a 'working draft' sequence. It currently consists of  
 13 contigs. The true order of the pieces is not known and their  
 order in this sequence record is arbitrary. Gaps between the  
 contigs are represented as runs N, but the exact sizes of the gaps  
 are unknown. This record will be updated with the finished sequence  
 as soon as it is available and the accession number will be  
 preserved

```

1      15418 contig of 15418 bp in length
15519      32589 contig of 17071 bp in length
32690      46533 contig of 13844 bp in length
46634      59641 contig of 13008 bp in length
59742      69969 contig of 10228 bp in length
70070      79400 contig of 9331 bp in length
79501      86689 contig of 7189 bp in length
86790      94162 contig of 7373 bp in length
94263      100141 contig of 5879 bp in length
100242      104375 contig of 4134 bp in length
104476      107052 contig of 2577 bp in length
107153      108671 contig of 1519 bp in length
108772      109149 contig of 378 bp in length

```

Sequence updated (01-Feb-2000)  
 Sequence updated (26-May-2000).

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 13 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

*      1      15418: contig of 15418 bp in length
*      15419      15518: gap of 100 bp
*      15519      32589: contig of 17071 bp in length
*      32590      32689: gap of 100 bp
*      32690      46533: contig of 13844 bp in length

```

```

*      46534      46633: gap of 100 bp
*      46634      59641: contig of 13008 bp in length
*      59642      59741: gap of 100 bp
*      59742      69969: contig of 10228 bp in length
*      69970      70069: gap of 100 bp
*      70070      79400: contig of 9331 bp in length
*      79401      79500: gap of 100 bp
*      79501      86689: contig of 7189 bp in length
*      86690      86789: gap of 100 bp
*      86790      94162: contig of 7373 bp in length
*      94163      94262: gap of 100 bp
*      94263      100141: contig of 5879 bp in length
*      100142      100241: gap of 100 bp
*      100242      104375: contig of 4134 bp in length
*      104376      104475: gap of 100 bp
*      104476      107052: contig of 2577 bp in length
*      107053      107152: gap of 100 bp
*      107153      108671: contig of 1519 bp in length
*      108672      108771: gap of 100 bp
*      108772      109149: contig of 378 bp in length.

```

## FEATURES

## source

```

1..109149
location/Qualifiers
  organism="Homo sapiens"
  mol_type="genomic DNA"
  db_xref="taxon:9606"
  chromosome="11"
  map="11q23"
  /clone="CMB9-2L13"
1..15418
  /note="assembly_fragment"
misc_feature
15519..32589
  /note="assembly_fragment"
misc_feature
32690..46533
  /note="assembly_fragment"
misc_feature
46634..59641
  /note="assembly_fragment"
misc_feature
59742..69969
  /note="assembly_fragment"
misc_feature
70070..79400
  /note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature
79501..86689
  /note="assembly_fragment"
misc_feature
86790..94162
  /note="assembly_fragment"
misc_feature
94263..100141
  /note="assembly_fragment"
misc_feature
100242..104375
  /note="assembly_fragment"
misc_feature
104476..107052
  /note="assembly_fragment"
misc_feature
107153..108671
  /note="assembly_fragment"
misc_feature
108772..109149
  /note="assembly_fragment clone_end:T7 vector_side:right"

```

## ORIGIN

## Query Match

Best Local Similarity 1.3%; Score 26; DB 2; Length 109149;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      2011 CCTACTAAATATATAAATTAGCTG 2036
      |||
Db      22347 CCTACTAAATATATAAATTAGCTG 22322

```

## RESULT 61

```

AC004898/c      AC004898      109867 bp      DNA      linear      PRI 30-SNP-2000
DEFINITION      Homo sapiens PAC clone RP4-814D15 from 7q35-q36, complete sequence.
ACCESSION      AC004898
VERSION      AC004898.3 GI:10440727
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)

```

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 109867)  
TITLE Suleiron,J.E., and Waterston,R.  
JOURNAL Toward a complete human genome sequence  
MEDLINE Genome Res. 8 (11), 1097-1108 (1998)  
99063792

PubMed  
PUBMED 9847074  
2 (bases 1 to 109867)  
AUTHORS Scott,K., Morales,R., Madsen,C. and Harper,M.  
REFERENCE The sequence of Homo sapiens PAC clone RP4-814D15  
JOURNAL Unpublished  
AUTHORS 3 (bases 1 to 109867)  
TITLE Waterston,R.H.  
JOURNAL Direct Submission  
AUTHORS Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 109867)  
REFERENCE Waterston,R.  
JOURNAL Direct Submission  
AUTHORS Submitted (30-SEP-2000) Department of Genetics, Washington  
JOURNAL University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Sep 30, 2000 this sequence version replaced gi:7630767.

COMMENT  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watscn.wustl.edu](mailto:sapiens@watscn.wustl.edu)  
----- Summary Statistics  
-----  
Center project name: H\_DJ0814D15  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://biopac.med.buffalo.edu>) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.  
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>) or from Pieter de Jong.  
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP5-1142J19, 200 base pair overlap. Actual start of this clone is at base position 65414 of RP5-1142J19; actual end is at base position 109867 of RP4-814D15.  
Location/Qualifiers  
1. 109867

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q35-q36"  
/clone="RP4-814D15"  
/clone\_lib="RPCI-4"  
1..24  
/note="similar to EST AA831457 (NID:g2904556) oc74h04.g1"  
1..24  
/note="similar to EST AW504069 (NID:g7141736)"  
1..24  
/note="similar to EST BE175498 (NID:g8638227)"  
568..860  
/rpt\_family="Alu"  
874..1099  
/rpt\_family="L1"  
1108..1416  
/rpt\_family="Alu"  
1428..1527  
/rpt\_family="L1"  
1528..1817  
/rpt\_family="Alu"  
1818..2479  
/rpt\_family="L1"  
2496..2589  
/note="similar to EST AW822656 (NID:g7915673) uq14a06.y1"  
2502..2589  
/note="similar to EST AA831457 (NID:g2904556) oc74h04.g1"  
2502..2589  
/note="similar to EST AW504069 (NID:g7141736)"  
2502..2589  
/note="similar to EST BE175498 (NID:g8638227)"  
2548..2589  
/note="similar to EST AA181620 (NID:g1765087) zp52d06.r1"  
2954..3054  
/rpt\_family="L1"  
3055..3349  
/rpt\_family="Alu"  
3350..3360  
/rpt\_family="L1"  
3361..3657  
/rpt\_family="Alu"  
3658..3852  
/rpt\_family="L1"  
4176..4706  
/rpt\_family="L1"  
4811..5113  
/rpt\_family="Alu"  
5240..5413  
/rpt\_family="Alu"  
5414..5722  
/rpt\_family="Alu"  
5533..5851  
/note="similar to EST AA091791 (NID:g1638318)"  
5605..5978  
/note="similar to EST AA180103 (NID:g1761368) zp15f01.r1"  
5827..5955  
/rpt\_family="Alu"  
6207..6418  
/note="similar to EST AA180348 (NID:g1761630) zp15f01.g1"  
6307..6411  
/note="similar to EST AA181620 (NID:g1765087) zp52d06.r1"  
6307..6411  
/note="similar to EST AW504069 (NID:g7141736)"  
6307..6411  
/note="similar to EST AW822656 (NID:g7915673) uq14a06.y1"  
6307..6402  
/note="similar to EST AA831457 (NID:g2904556) oc74h04.g1"  
6307..6372  
/note="similar to EST BE175498 (NID:g8638227)"  
6308..6411  
/note="similar to EST AW390729 (NID:g6895388)"

```
repeat_region      6694..6924
                    /rpt_family="L1"
misc_feature        8002..8294
                    /note="similar to EST BE008943 (NID:g8269176)"
repeat_region      9366..9478
                    /rpt_family="MER53"
repeat_region      9812..9983
                    /rpt_family="MER1_type"
misc_feature        9946..10236
                    /note="similar to EST AA582469 (NID:g2359829) m54f04.s1"
repeat_region      10222..10523
                    /rpt_family="Alu"
misc_feature        11064..11451
                    /note="similar to EST AA630532 (NID:g2903631) cc52c04.s1"
misc_feature        11208..11811
                    /note="similar to EST AW73446 (NID:g8164517)"
repeat_region      11645..11768
                    /rpt_family="T2_type"
repeat_region      12707..12805
                    /rpt_family="MIR"
repeat_region      13094..13378
                    /rpt_family="Alu"
misc_feature        13782..14271
                    /note="similar to EST AW446874 (NID:g6988661)"
misc_feature        13858..14422
                    /note="similar to EST AA522017 (NID:g2262762) vH78e09.r1"
repeat_region      14903..15215
                    /rpt_family="Alu"
misc_feature        15243..15440
                    /note="similar to EST BE158233 (NID:g8620954)"
misc_feature        15268..15575
                    /note="similar to EST AI064719 (NID:g6358991)"
misc_feature        15343..15577
                    /note="similar to EST AA181620 (NID:g1765087) zp52d06.r1"
misc_feature        15343..15575
                    /note="similar to EST AW390729 (NID:g6895388)"
misc_feature        15343..15575
                    /note="similar to EST AW822656 (NID:g7915673) uq14a06.y1"
misc_feature        15343..15397
                    /note="similar to EST AW504069 (NID:g7141736)"
misc_feature        15529..15575
                    /note="similar to EST AW907921 (NID:g8073150) ur78a07.y1"
misc_feature        15723..15880
                    /note="similar to EST AI064719 (NID:g6358991)"
misc_feature        15723..15880
                    /note="similar to EST AW907921 (NID:g8073150) ur87a07.y1"
misc_feature        15723..15844
                    /note="similar to EST AW390729 (NID:g6895388)"

Query Match      1.3%; Score 26; DB 9; Length 109867;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2011 CCTACTAAATATTAATAATAGCTG 2036
        |||||
Db      49084 CCTACTAAATATTAATAATAGCTG 49059
```

```
RESULT 62
AL583823_0/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL583823 Accession AL583823
Fragment Name      Begin      End
AL583823_1         100001    210000
AL583823_2         200001    310000
AL583823_3         300001    360756
LOCUS      AL583823
DEFINITION Homo sapiens chromosome 9 clone RP11-180F6, 26 unordered pieces.
ACCESSION  AL583823
VERSION     AL583823.6 GI:13591508
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
```

```
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        Plumb,B.
JOURNAL      Direct Submission
             Submitted (08-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
             requests: clonerequest@sanger.ac.uk
             On Apr 10, 2001 this sequence version replaced gi:13568269.
COMMENT      ----- Genome Center
             Center: Sanger Centre
             Center code: SC
             Web site: http://www.sanger.ac.uk
             Contact: humquery@sanger.ac.uk
             ----- Project Information
             Center project name: BA180F6
             ----- Summary Statistics
             Assembly program: XGAP4; version 4.5
             Sequencing vector: plasmid; 108752; 100% of reads
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Consensus quality: 353030 bases at least Q40
             Consensus quality: 355077 bases at least Q30
             Consensus quality: 356660 bases at least Q20
             Insert size: 358256; sum-of-contigs
             Insert size: 334181; 3.6% error; agarose-fp
             Quality coverage: 9.14x in Q20 bases; sum-of-contigs Quality
             coverage: 10.15x in Q20 bases; agarose-fp
             -----
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 26 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             1
             5322: contig of 5322 bp in length
             5422: gap of 100 bp
             5423
             15390: contig of 9968 bp in length
             15391
             15490: gap of 100 bp
             15491
             46218: contig of 30728 bp in length
             46219
             46318: gap of 100 bp
             46319
             58017: contig of 11699 bp in length
             58018
             58117: gap of 100 bp
             58118
             80558: contig of 2241 bp in length
             80559
             80558: gap of 100 bp
             80559
             100032: contig of 19374 bp in length
             100033
             100132: gap of 100 bp
             100133
             108996: contig of 8864 bp in length
             108997
             109096: gap of 100 bp
             109097
             111845: contig of 279 bp in length
             111846
             111945: gap of 100 bp
             111946
             127953: contig of 16008 bp in length
             127954
             128053: gap of 100 bp
             128054
             147475: contig of 19422 bp in length
             147476
             147575: gap of 100 bp
             147576
             159043: contig of 11468 bp in length
             159044
             159143: gap of 100 bp
             159144
             161504: contig of 2361 bp in length
             161505
             161604: gap of 100 bp
             161605
             169773: contig of 8169 bp in length
             169774
             169873: contig of 8169 bp in length
             169874
             178082: contig of 8209 bp in length
             178083
             178182: gap of 100 bp
             178183
             185590: contig of 7408 bp in length
             185591
             185690: gap of 100 bp
             185691
             200490: contig of 14800 bp in length
             200491
             200590: gap of 100 bp
             200591
             213368: contig of 12678 bp in length
             213369
             213547: contig of 2179 bp in length
             213548
             215647: gap of 100 bp
```



```

* 215648 24011: contig of 24464 bp in length
* 24012 24021: gap of 100 bp
* 240212 245870: contig of 5659 bp in length
* 245871 245970: gap of 100 bp
* 245971 251519: contig of 5549 bp in length
* 251520 251619: gap of 100 bp
* 251620 261555: contig of 9936 bp in length
* 261556 261656: gap of 100 bp
* 261657 306526: contig of 44871 bp in length
* 306527 306627: gap of 100 bp
* 306628 316314: contig of 9688 bp in length
* 316315 316414: gap of 100 bp
* 316415 350996: contig of 34582 bp in length
* 350997 351096: gap of 100 bp
* 351097 360756: contig of 9660 bp in length.
FEATURES
SOURCE
1..360756
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-180F6"
/clone_11b="RPC1-11.1"
1..5322
/note="assembly_fragment:00068
fragment_chain:1"
misc_feature
5423..15390
/note="assembly_fragment:02309
fragment_chain:1"
misc_feature
15491..46218
/note="assembly_fragment:04749
fragment_chain:1"
misc_feature
46319..58017
/note="assembly_fragment:06120
fragment_chain:1"
misc_feature
58118..80558
/note="assembly_fragment:07209
fragment_chain:1"
misc_feature
80659..100032
/note="assembly_fragment:00462
fragment_chain:2"
misc_feature
100133..108996
/note="assembly_fragment:02942
fragment_chain:2"
misc_feature
109097..111845
/note="assembly_fragment:07600
fragment_chain:2"
misc_feature
111946..127953
/note="assembly_fragment:02410
fragment_chain:2"
misc_feature
128054..147475
/note="assembly_fragment:04056
fragment_chain:3"
misc_feature
147576..159043
/note="assembly_fragment:01489
fragment_chain:3"
misc_feature
159144..161504
/note="assembly_fragment:05231
fragment_chain:3"
misc_feature
161605..169773
/note="assembly_fragment:04982
fragment_chain:4"
misc_feature
169874..178082
/note="assembly_fragment:04824
fragment_chain:4"
misc_feature
178183..185590
/note="assembly_fragment:08296
fragment_chain:4"
misc_feature
185691..200490
/note="assembly_fragment:06936
fragment_chain:5"
misc_feature
200591..213268
/note="assembly_fragment:01198

```

```

fragment_chain:5"
21369..215547
/note="assembly_fragment:02963
fragment_chain:5"
misc_feature
215648..240111
/note="assembly_fragment:04139
fragment_chain:6"
misc_feature
240212..245870
/note="assembly_fragment:03806
fragment_chain:6"
misc_feature
245971..251519
/note="assembly_fragment:05149
fragment_chain:7"
misc_feature
251620..261555
/note="assembly_fragment:02527
fragment_chain:7"
misc_feature
261656..306526
/note="assembly_fragment:03577"
306627..316314
/note="assembly_fragment:03882"
316415..350996
/note="assembly_fragment:04545"
351097..360756
/note="assembly_fragment:07734"
ORIGIN
Query Match 1.3%; Score 26; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2011 CCTACTAAATATATAATATAGCTG 2036
Db 24518 CCTACTAAATATATAATATAGCTG 24493
RESULT 63
AF030876 112752 bp DNA linear PRI 24-OCT-2002
LOCUS
DEFINITION
Homo sapiens chromosome X clone Qc-8D3, RP4-671D9, 1104-219D.
L110-C1837 map q28, complete sequence.
ACCESSION
AF030876 AF031075 AF031076 AF031077 AF031078
VERSION
AF030876.2 GI:22830571
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 112752)
Reichwald K., Thiesen J., Wiehe T., Weitzel J., Poustka W.A.,
Roenthal A., Platzer M., Strätling W.H. and Kloeckle P.
Comparative sequence analysis of the MCCP2-locus in human and mouse
reveals new transcribed regions
Mamm. Genome 11 (3), 182-190 (2000)
TITLE
JOURNAL
MEDLINE
20188769
PUBMED
10723722
2 (bases 1 to 112752)
Reichwald K. and Platzer M.
Direct Submission
Submitted (21-OCT-1997) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 112752)
Reichwald K. and Platzer M.
Direct Submission
Submitted (13-SEP-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thüringia 07745, Germany
On or before Sep 13, 2002 this sequence version replaced
gi:6649930, gi:6649933, gi:6649934, gi:3002592, gi:3002589.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information

```

```
Center project name: x8+
Center clone name: Qc-8D3, RP4-671D9, I104-219D, L110-C1837
----- Summary Statistics -----
Sequencing vector: M13mp18: 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 112735 bases at least Q40
Consensus quality: 112752 bases at least Q30
Consensus quality: 112752 bases at least Q20
Quality coverage: 15.83x
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data; an attempt was made to
resolve all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone or more
than one M13 subclone; and the assembly was confirmed by
restriction digest.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates. Base-by-base quality values are not generally
visible from the GenBank flat file format but are available
as part of this entry's ASN.1 file.
-----
FEATURES
source      Location/Qualifiers
            1..112752
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="X"
             /map="q28"
             /clone="Qc-8D3, RP4-671D9, I104-219D, L110-C1837"
source      1..44173
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="Qc-8D3"
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="RP4-671D9"
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="1104-219D"
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="1104-219D"
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="1110-C1837"
             /clone="L110-C1837"
             /clone="11983.11363"
             /gene="IRAK1"
             /evidence=experimental
             /complement(join(1983..3389,4001..4150,4515..4905,
5514..5750,7504..7569,7909..8116,8421..8539,9478..9592,
9763..9827,10071..10259,10447..10550,10669..10800,
10903..11070,11149..11363))
             /product="Interleukin-1 receptor-associated kinase 1 mRNA"
             /evidence=experimental
             /complement(1993..1998)
             /gene="IRAK1"
             /evidence=experimental
             /complement(join(331..3389,4001..4150,4515..4905,
5514..5750,7504..7569,7909..8116,8421..8539,9478..9592,
9763..9827,10071..10259,10447..10550,10669..10800,
10903..11070,11149..11284))
             /gene="IRAK1"
             /codon_start=1
             /evidence=experimental
             /product="Interleukin 1 receptor-associated kinase 1"
polyA_signal
```

```
/protein_id="AAC08756.1"
/db_xref="gi:3002590"
/translation="MAGSGPGCEPPAPAPGAQHFLEYVPPMMWCRFYKWDALPEADMCQ
FAALIVPOTETLRCERSGORTASVLMWIRNNAVALVHILHTLQLLRADITLAW
HPRALVPSPGTPAPRSSIRAPAEAEASPKRLPSASTFLSPARGSCOTHSPEELKI
VPSFSLMPPEPSAPSSTKPEPSSVLLQARFPFCWICETISRGTHNSEELKI
GEGGQCVYRAVMNTVYAVRKLKNALDEWVAVQSLVTEQSLSRHPNIVDFAG
YCAQNGFYCLVYGLPNGLSLEDRLLHCOTQACPLSPQRLDILGTARAIQPLQDPS
SLIHGDIKSNVILDERLTPKLDGRLAFSPASSPSQSMVARTOTVGTLLYP
BEYIKTGRILAVDTPTEFVVVLETLAQORAVKTHGARVKYLKDLVEEAEAEVALR
SQSTQLAGLAADANAAPIAMQIYKXHLDPKRGPCPEPLGIGQLACCCIRAPAKR
PMTQVTVRLEKQLQVAVGVRHNSAASCTPPSPENSVYSTGASHGAPWQPLA
PSGAQAQAEQLQKRGPNQPVSDSLGSLALRSMWLTSPCPLDAPLREXGCGQD
TAGESWMSGGSGSRPTAVEGALSSASSSSPQIITINPARKWQKIALYEDGALD
SLQLSSSLPQLGLEQDRQGPESSEDFQS"
10643..11676
/note="CPG island: 77.1 %GC, o/e=0.86, #CPGs=129"
10703..11012
/note="single stranded/single chemistry region"
11526..11650
/note="single stranded/single chemistry region"
complement(113285..89150)
/gene="MECP2"
/evidence=experimental
complement(113304..11309)
/gene="MECP2"
/evidence=experimental
20563
/note="C substituted in clone: RP4-671D9"
/replace="C"
complement(join(21712..22922,23679..24029,83663..83786,
89082..89150))
/gene="MECP2"
/product="Methyl-CPG binding protein 2 mRNA, short 3'UTR"
/evidence=experimental
complement(21731..21736)
/gene="MECP2"
/evidence=experimental
complement(join(21839..22922,23679..24029,83663..83688))
/gene="MECP2"
/codon_start=1
/evidence=experimental
/product="Methyl-CPG binding protein 2"
/protein_id="AAC08757.1"
/db_xref="gi:3002591"
/translation="MVAGMLGLREKSESDQLQGLKRLPLFKYVKDKKEKEGKHE
PVQPSAHNSAPAEAGKAETSGSGSIIPKRSRHSIIDRPQMVDDPFLPE
GWTRLKLRKSGRSAGKYDVYLINQGAFAFSKELAYFEKVGDTSLDPNDFDTLV
GKGSFRRQKRPKPKPSKPAKGTGKGRFPGSSTTPKATSGVOVKVLEKSPG
KLLVMPFQTSPPGAEAGGATSTQVWVIRPKGRKRALEDPOAIPIKRRGKPSV
AAAALFAKKAVKESISVQETVLPKIKRTRETVSJEVEVVPDLVLTSGESGK
GLTKTSKPSKESSPKRSASPSPKENHNHNSBPKAVPLPLPPEPK
PSSSDPSPPEPDLSGVCKEKMPRGSLSDGCKEPAKTPAVATATAEKY
KIRGGEKDIVSSMPRNPBEPTSTPTERS"
26018
/note="A substituted in clone: RP4-671D9"
/replace="T"
27674
/note="C substituted in clone: I104-219D"
/replace="A"
28639..28658
/note="deleted in clone: I104-219D"
/replace=""
28655..28658
/note="deleted in clone: Qc-8D3"
/replace=""
28669..28670
/note="deleted in clone: Qc-8D3, deleted in clone:
```

RP4-671D9"  
/replace=""  
34248  
/note="G substituted in clone: 1104-219D"  
/replace="A"  
35662  
/note="deleted in clone: Qc-8D3 , deleted in clone:  
RP4-671D9"  
/replace=""  
37415 .37418  
/note="deleted in clone: 1104-219D"  
/replace=""  
37417 .37418  
/note="deleted in clone: Qc-8D3 , deleted in clone:  
RP4-671D9"  
/replace=""  
37429  
/note="deleted in clone: 1104-219D , deleted in clone:  
37429

Query Match 1.3%; Score 26; DB 9; Length 112752;  
Best Local Similarity 100.0%; Pred.No. 0.043;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2011 CCTACTAAATATATAATAGCTG 2036  
|||||  
Db 98876 CCTACTAAATATATAATAGCTG 98901

RESULT 64  
AC010331 114491 bp DNA 1linear PRI 13-JUL-2002  
LOCUS Home sapiens chromosome 19 clone CTD-3001H11, complete sequence.  
DEFINITION AC010331  
AC010331.8 GI:21743758  
VERSION HTG.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 114491)  
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
UNPUBLISHED  
2 (bases 1 to 114491)  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 114491)  
REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (28-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 114491)  
REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Jul 13, 2002 this sequence version replaced gi:1579579.  
COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.3.  
Location/Qualifiers  
1. 114491  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-3001H11"

ORIGIN  
Query Match 1.3%; Score 26; DB 9; Length 112752;

Query Match 1.3%; Score 26; DB 9; Length 114491;  
Best Local Similarity 100.0%; Pred.No. 0.043;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2011 CCTACTAAATATATAATAGCTG 2036  
|||||  
Db 39839 CCTACTAAATATATAATAGCTG 39864

RESULT 65  
AC008102 126787 bp DNA 1linear PRI 09-APR-2003  
LOCUS Homo sapiens chromosome 11 clone bac11121 map 11q13, complete  
sequence.  
DEFINITION AC008102  
AC008102 GI:29650247  
VERSION HTG.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 126787)  
TITLE Lewis,J., Smith,D., Srivatsan,E. and Roe,B.A.  
JOURNAL Direct Submission  
UNPUBLISHED  
2 (bases 1 to 126787)  
REFERENCE  
AUTHORS Lewis,J., Smith,D., Srivatsan,E. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (22-JUL-1999) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
3 (bases 1 to 126787)  
REFERENCE  
AUTHORS Lewis,J., Smith,D., Srivatsan,E. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (28-FEB-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
4 (bases 1 to 126787)  
REFERENCE  
AUTHORS Lewis,J., Smith,D., Srivatsan,E. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (29-FEB-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
5 (bases 1 to 126787)  
REFERENCE  
AUTHORS Lewis,J., Smith,D., Srivatsan,E. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (12-AUG-2000) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
6 (bases 1 to 126787)  
REFERENCE  
AUTHORS Lewis,J., Smith,D., Srivatsan,E. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (09-APR-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
On Apr 9, 2003 this sequence version replaced gi:9797815.  
COMMENT ----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR  
-----  
Location/Qualifiers  
1. 126787  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13"  
/clone="bac41121"

ORIGIN  
Query Match 1.3%; Score 26; DB 9; Length 126787;

Best Local Similarity 100.0%; Pred. No. 0.042;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2011 CCTACTAAATATATAATAGCTG 2036  
Db 105961 CCTACTAAATATATAATAGCTG 105986

RESULT 66  
AL136458/c 128726 bp DNA linear PRI 08-FEB-2002  
LOCUS Human DNA sequence from clone RP4-537K17 on chromosome 1p32.1-32.3,  
DEFINITION complete sequence.  
ACCESSION AL136458  
VERSION AL136458.16 GI:18643738  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 128726)  
AUTHORS Frankland, J.  
TITLE Direct Submission  
JOURNAL Submitted (08-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk  
On Feb 10, 2002 this sequence version replaced gi:18643738.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Emi, EMBL; Swi,  
SWISSPROT; Trl, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
This sequence was generated from part of bacterial clone contigs of human  
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr1  
RP4-537K17 is from the library RPCI-4 constructed by the group of  
Pleier de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2

This sequence is the entire insert of clone RP4-537K17 The true  
left end of clone RP5-1155K23 is at 4366 in this sequence. The  
true right end of clone RP11-112D15 is at 22992 in this sequence.  
Location/Qualifiers

FEATURES  
source 1..128726  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="p32.1-32.3"  
/clone="RP4-537K17"  
/clone\_1lb="RPCI-4"  
18903..18957  
misc\_feature  
/note="Sequence from overlapping clone RP11-112D15  
(AL590374). Assembly confirmed by restriction digest."  
95699  
misc\_feature  
/note="Tandem repeat. Forced join. Gap size estimated to  
be approximately 500bp by restriction digest data."  
98263  
misc\_feature  
/note="Sequence from overlapping clone RP5-1155K23

(AL162739). Assembly confirmed by restriction digest."  
misc\_feature 112298..112395  
/note="Sequence from overlapping clone RP5-1155K23  
(AL162739). Assembly confirmed by restriction digest."  
misc\_feature 124561..124562  
/note="Sequence from overlapping clone RP5-1155K23  
(AL162739). Assembly confirmed by restriction digest."  
ORIGIN

Query Match 1.3%; Score 26; DB 9; Length 128726;  
Best Local Similarity 100.0%; Pred. No. 0.042;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2011 CCTACTAAATATATAATAGCTG 2036  
Db 9575 CCTACTAAATATATAATAGCTG 9550

RESULT 67  
AC090170 130195 bp DNA linear HTG 26-AUG-2001  
LOCUS Homo sapiens chromosome 8 clone RP4-537K17 map 8, WORKING DRAFT  
DEFINITION SEQUENCE, 3 ordered pieces.  
ACCESSION AC090170 GI:15290842  
VERSION AC090170.4  
KEYWORDS HTG; HTGS PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 130195)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
JOURNAL Unpublished  
TITLE Homo sapiens chromosome 8, clone RP4-537K17  
REFERENCE 2 (bases 1 to 130195)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Baetsen, V., Boguslavsky, L., Boukhalter, B., Brown, A.,  
Cammarata, V., Campopiano, A., Choepel, T., Colangelo, M., Collins, S.,  
Collumore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,  
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,  
Lohoczky, D., Levine, R., Liu, G., Maclean, C., MacDonald, P.,  
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,  
McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunhkhang, P., Piere, N., Pollara, V., Raymond, C., Retta, R.,  
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,  
Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
Souarez, N., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 26, 2001 this sequence version replaced gi:14717342.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/BM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: W1BR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: 142117  
Center clone name: 537\_K17  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads

```

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 128904 bases at least Q40
Consensus quality: 129439 bases at least Q30
Consensus quality: 129708 bases at least Q20
Insert size: 127000; agarose-fp
Insert size: 129995; sum-of-ctrls
Quality coverage: 8.4 in Q20 bases; agarose-fp
Quality coverage: 8.2 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 59629: contig of 59629 bp in length
* 59630 59729: gap of 100 bp
* 59730 93246: contig of 33517 bp in length
* 93247 93346: gap of 100 bp
* 93347 130195: contig of 36849 bp in length.
*
Location/Qualifiers
1..130195
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP4-537K17"
/clone_lib="RPCT Human PAC library 4"
1..59629
/notes="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
59730..93246
/notes="assembly_fragment"
93347..130195
/notes="assembly_fragment"
clone_end:T7
vector_side:right"

ORIGIN

Query Match 1.3%; Score 26; DB 2; Length 130195;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036
|||||
Db 120626 CCTACTAAATATATAAATTAGCTG 120651

RESULT 68
AC008940 131975 bp DNA linear PRI 18-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2319M24, complete sequence.
AC008940
VERSION AC008940.3 GI:7381726
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 131975)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 131975)
REFERENCE DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
JOURNAL

```

```

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 131975)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 131975)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (18-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Apr 1, 2000 this sequence version replaced gi:6165127.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.3.
STS Content:
SHGC-79029 G48806.
Location/Qualifiers
1..131975
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2319M24"

ORIGIN

Query Match 1.3%; Score 26; DB 9; Length 131975;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036
|||||
Db 120692 CCTACTAAATATATAAATTAGCTG 120717

RESULT 69
AC012435 132994 bp DNA linear PRI 01-NOV-2002
LOCUS Homo sapiens chromosome 15, clone RP11-10017, complete sequence.
AC012435
VERSION AC012435.13 GI:24462398
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 132994)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 132994)
Birken, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barina, N., Becker, R., Bogdanovskiy, L., Boukhalter, B.,
Brown, A., Casle, A., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., DeArnell, K., Dewar, K., Domingo, M., Donnell, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
Lehoczky, I., Lien, C., Locke, K., Macdonald, P., Margulis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., McDermott, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testave, S., Tjirell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
JOURNAL Submitted (27-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 132994)
Birken, B., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,

```

Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatae,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Mathews,C., McCarthy,M., Meldrum,J., Meneus,L., Mhoya,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Tophan,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmerer,A. and Zody,M.

Direct Submission  
Submitted (13-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 132994)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Baran,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatae,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Mathews,C., McCarthy,M., Meldrum,J., Meneus,L., Mhoya,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Tophan,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmerer,A. and Zody,M.

Direct Submission  
Submitted (01-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 1, 2002 this sequence version replaced gt:23928489.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE  
JOURNAL  
COMMENT

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3124  
Center clone name: 10\_O\_17  
-----

Only the first 133,0 Kilobases of this clone are being submitted.  
The remainder overlaps accession number AC100835 [MIGR project L21559].

FEATURES  
source  
1. 132994  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP11-10017"  
/clone\_1lb="RP11-10017 Human Male BAC"  
complement(1010..1134)  
/rpt\_family="MIR3"  
complement(5319..5502)  
/rpt\_family="MIR"  
complement(5563..5811)  
repeat\_region  
repeat\_region  
repeat\_region  
repeat\_region

/rpt\_family="MIR3"  
6015..6040  
/rpt\_family="TCCC)n"  
9276..9363  
/rpt\_family="(CA)n"  
complement(11379..11499)  
/rpt\_family="MIR"  
complement(12548..12709)  
/rpt\_family="MIR3"  
complement(13629..13633)  
/note="<30 qual SNGL region"  
complement(18206..18451)  
/rpt\_family="MIR"  
complement(18445..18538)  
/rpt\_family="MIR3"  
complement(18639..18689)  
/rpt\_family="MIR54A"  
18990..19298  
/rpt\_family="AluSx"  
complement(19299..19365)  
/rpt\_family="MIR54A"  
19625..19783  
/rpt\_family="FRAM"  
19984..20134  
/rpt\_family="L2"  
complement(20435..20614)  
/rpt\_family="AluSx"  
21312..21497  
/rpt\_family="MIR3"  
22855..23152  
/rpt\_family="L2"  
complement(23365..23520)  
/rpt\_family="MIR3"  
26208..26240  
/rpt\_family="(CGGGG)n"  
26808..26840  
/rpt\_family="(CAG)n"  
complement(28724..28825)  
/rpt\_family="MIR"  
complement(28831..29054)  
/rpt\_family="MIR"  
complement(30150..30309)  
/rpt\_family="L3"  
complement(30358..30498)  
/rpt\_family="Charlie8"  
30499..30812  
/rpt\_family="AluSg"  
30816..30887  
/rpt\_family="AluSg"  
30927..31160  
/rpt\_family="AluSg"  
32167..32267  
/rpt\_family="MIR3"  
32385..32665  
/rpt\_family="AluSg"  
32666..32695  
/rpt\_family="(CTAA)n"  
complement(32852..33264)  
/rpt\_family="L3"  
complement(33305..33418)  
/rpt\_family="L3"  
33487..33578  
/rpt\_family="MER117"  
33623..33919  
/rpt\_family="AluY"  
34332..34448  
/rpt\_family="MIR"  
35997..36141  
/rpt\_family="MIR"  
36190..36442  
/rpt\_family="L2"  
36533..36812  
/rpt\_family="L2"

```

repeat_region 36817..37123
                /rpt_family="AluSx"
repeat_region 37755..38156
                /rpt_family="L3"
repeat_region 40420..40703
                /rpt_family="AluUo"
repeat_region 40704..40728
                /rpt_family="AluUo"
repeat_region 40945..41214
                /rpt_family="CMATn"
repeat_region 43370..43498
                /rpt_family="L2"
repeat_region 43370..43498
                /rpt_family="MIR3"
repeat_region 46010
                /rpt_family="AluSx"
repeat_region 46633
                /rpt_family="L2"
repeat_region 47260
                /rpt_family="MIR"
repeat_region 47402
                /rpt_family="MER90"
repeat_region 47696
                /rpt_family="MER90"
                Complement(47403..47696)

Query Match 1.3%; Score 26; DB 9; Length 132994;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATTTAAATTTAGCTG 2036
DB 32499 CCTACTAAATTTAAATTTAGCTG 32524

RESULT 70
AC090615/c 142895 bp DNA linear PRI 02-JUL-2002
LOCUS Homo sapiens chromosome 17, clone RP11-578C11, complete sequence.
AC090615
AC090615.9 GI:21672217
HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 142895)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-578C11
2 (bases 1 to 142895)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Baetien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Cammarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczyk,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Menus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Travers,M., Travls,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimer,A. and Zody,M.
Direct Submission
Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 142895)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

```

```

Anderson,S., Barna,N., Baetien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Cammarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glnde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Karatas,A., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Roh,K.,
Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimer,A. and Zody,M.
Direct Submission
Submitted (10-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 142895)
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Cammarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A.,
Landers,T., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimer,A. and Zody,M.
Direct Submission
Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 2, 2002 this sequence version replaced gi:21326276.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12035
Center clone name: 578_C_11
-----
FEATURES
source
1..142895
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-578C11"
/clone_lib="RP11 Human Male BAC"
531..598
/rpt_family="MIR"
608..684
/rpt_family="AluUo"
1059..1208
repeat_region

```

```

/rpt_family="L2"
repeat_region 1231.1544
/rpt_family="L2"
repeat_region 1545.1856
/rpt_family="AluSg"
repeat_region 1857.2249
/rpt_family="L2"
repeat_region 2250.2550
/rpt_family="AluSx"
repeat_region 2551.2599
/rpt_family="L2"
repeat_region 2603.2644
/rpt_family="GGGA)n"
repeat_region 3033.3033
/rpt_family="LFR33A"
repeat_region 3121.3205
/rpt_family="L2"
repeat_region 3207.3519
/rpt_family="AluUb"
repeat_region 3830.3999
/rpt_family="MIR"
repeat_region 4734.4834
/rpt_family="CT-rich"
repeat_region 5000.5000
/rpt_family="L2"
repeat_region 5086.5164
/rpt_family="L2"
repeat_region 5798.6092
/rpt_family="AluUb"
repeat_region 6101.6378
/rpt_family="AluUb"
repeat_region 6707.7028
/rpt_family="AluSg"
repeat_region 7361.7534
/rpt_family="MER45A"
repeat_region 7946.8102
/rpt_family="MIR"
repeat_region 8104.8508
/rpt_family="L2"
repeat_region 8790.8790
/rpt_family="MER47A"
repeat_region 9021.9021
/rpt_family="AluY"
repeat_region 9022.9055
/rpt_family="(TTG)n"
repeat_region 9358.9358
/rpt_family="AluSg"
repeat_region 9475.9475
/rpt_family="MER47A"
repeat_region 9509.9581
/rpt_family="LIM4"
repeat_region 9584.9895
/rpt_family="AluSx"
repeat_region 9910.10217
/rpt_family="AluSg"
repeat_region 10232.10317
/rpt_family="GA-rich"
repeat_region 10316.10630
/rpt_family="AluUb"
repeat_region 10631.10725
/rpt_family="GA-rich"
repeat_region 10726.10860
/rpt_family="MER89"
repeat_region 11008.11008
/rpt_family="L2"
repeat_region 11009.11326
/rpt_family="AluSx"
repeat_region 11327.11632
/rpt_family="L2"
repeat_region 11675.11987
/rpt_family="AluUb"
repeat_region 12066.12147
/rpt_family="L2"

```

```

repeat_region complement(12775.13106)
/rpt_family="AluSx"
repeat_region complement(1331.13364)
/rpt_family="AluUb"
repeat_region complement(13368.13691)
/rpt_family="LIMC/D"
repeat_region complement(13793.14098)
/rpt_family="AluUb"
repeat_region complement(14168.14442)
/rpt_family="AluUb"
repeat_region complement(14455.14755)
/rpt_family="AluSg"
repeat_region complement(14838.15140)
/rpt_family="AluSg"

```

Query Match 1.3% Score 26; DB 9; Length 142895;  
 Best Local Similarity 100.0%; Pred.No. 0.041;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2011 CCTACTAAATATATAATTAGCTG 2036  
 Db 78739 CCTACTAAATATATAATTAGCTG 78714

RESULT 71  
 AC092574/c 144117 bp DNA linear PRI 01-MAR-2002  
 LOCUS  
 DEFINITION Homo sapiens BAC clone RP11-2H3 from 4, complete sequence.  
 ACCESSION AC092574 AC015776  
 VERSION AC092574.2 GI:15638708  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Sulston,J.E. and Waterston,R.  
 TITLE Toward a complete human genome sequence  
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
 MEDLINE 99063792  
 PUBMED 9847074

REFERENCE  
 AUTHORS Haglund,K., Haakenson,W. and Spalding,L.  
 TITLE The sequence of Homo sapiens BAC clone RP11-2H3  
 JOURNAL Unpublished (2001)  
 REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JUN-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 COMMENT 5 (bases 1 to 144117)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 COMMENT On Sep 18, 2001 this sequence version replaced gi:14916159.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc  
 Contact: saplens@watsn.wustl.edu  
 ----- Summary Statistics  
 Center project name: H\_NH0002H03  
 Drafting Center: WIBR  
 -----



NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateo, M., Caranese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-478C6, 2000 bp overlap. Actual end is at base position 144117 of RP11-2H3.

Data from AC079140 was used to finish this clone, AC092574.

Sequence derived from a single plasmid subclone from base position 24830 to 24840.

The sequence of AC015776 has been incorporated into AC092574.

FEATURES  
source  
location/Qualifiers  
1..14417  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"  
/clone="RP11-2H3"  
/clone\_lib="RPCT-11"  
1..51  
/rpt\_family="AT-rich"  
repeat\_region  
138..426  
/rpt\_family="Alu"  
misc\_feature  
138..159  
/note="similar to Homo sapiens EST A1911406 (NID:g55631142)  
wd16g05.x1"  
misc\_feature  
362..421  
/note="similar to Homo sapiens EST BG914267  
(NID:g14294743)"  
misc\_feature  
378..427  
/note="similar to Homo sapiens EST BE159632  
(NID:g8622353)"  
misc\_feature  
393..426  
/note="similar to Homo sapiens EST BF797755  
(NID:g12102809)"  
repeat\_region  
464..738  
/rpt\_family="Alu"  
repeat\_region  
1116..1481  
/rpt\_family="L1"  
misc\_feature  
1195..1200  
/note="similar to Homo sapiens EST BE842803  
(NID:g10275181)"  
misc\_feature  
1371..1373  
/note="similar to Homo sapiens EST AW044288 (NID:g5904817)

misc\_feature  
1552..1649  
/note="similar to Homo sapiens EST BF675662  
(NID:g11949557)"  
misc\_feature  
1563..1919  
/note="similar to Homo sapiens EST BE883118  
(NID:g10331894)"  
misc\_feature  
1564..1581  
/note="similar to Homo sapiens EST BG721813  
(NID:g14001000)"  
misc\_feature  
1614..1970  
/note="similar to Homo sapiens EST BG391074  
(NID:g13284522)"  
misc\_feature  
1614..1844  
/note="similar to Homo sapiens EST BE177799  
(NID:g8656951)"  
misc\_feature  
1614..1733  
/note="similar to Homo sapiens EST BE781790  
(NID:g10202988)"  
misc\_feature  
1614..1733  
/note="similar to Homo sapiens EST BG754870  
(NID:g14065523)"  
misc\_feature  
1620..1859  
/note="similar to Homo sapiens EST BG289439  
(NID:g13045285)"  
misc\_feature  
1622..1854  
/note="similar to Homo sapiens EST AL119535  
(NID:g5925434)"  
misc\_feature  
1622..1849  
/note="similar to Homo sapiens EST AW409714 (NID:g6935255)  
fh02a02.x1"  
misc\_feature  
1622..1746  
/note="similar to Homo sapiens EST AW576720  
(NID:g7251860)"  
misc\_feature  
1636..1841  
/note="similar to Homo sapiens EST BG721813  
(NID:g14001000)"  
misc\_feature  
1644..1879  
/note="similar to Homo sapiens EST AU571293  
(NID:g12928444)"  
misc\_feature  
1673..2021  
/note="similar to Homo sapiens EST BE172609  
(NID:g8635335)"  
misc\_feature  
1676..1901  
/note="similar to Homo sapiens EST BE694258  
(NID:g10081418)"  
misc\_feature  
1686..1901  
/note="similar to Homo sapiens EST BE694276  
(NID:g10081436)"  
misc\_feature  
1760  
/note="similar to Homo sapiens EST AA614249 (NID:g2466383)  
np09g09.s1"  
misc\_feature  
1793..1865  
/note="similar to Homo sapiens EST BG754870  
(NID:g14065523)"  
misc\_feature  
1793..1854  
/note="similar to Homo sapiens EST BE781790  
(NID:g10202988)"  
misc\_feature  
1810..1909  
/note="similar to Homo sapiens EST BG616121  
(NID:g13667492)"  
misc\_feature  
1813..1970  
/note="similar to Homo sapiens EST AA397793 (NID:g2051001)  
zt88e05.x1"  
misc\_feature  
1826..1875  
/note="similar to Homo sapiens EST AU133219  
(NID:g10993758)"  
misc\_feature  
1831..1932  
/note="similar to Homo sapiens EST BG287904  
(NID:g13042206)"  
misc\_feature  
1831..1932  
/note="similar to Homo sapiens EST BG741860  
(NID:g14052513)"

```

misc_feature      1850..1865
                  /note="similar to Homo sapiens EST AA614249 (NID:g246383)
                  np09g09.g1"
misc_feature      1851..1870
                  /note="similar to Homo sapiens EST BE826200
                  (NID:g10258578)"
misc_feature      1852..2015

Query Match      1.3% Score 26; DB 9; Length 144117;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2011 CCTACTAAATATATAAATAGCTG 2036
Db      53725 CCTACTAAATATATAAATAGCTG 53700

RESULT 72
AC117469      144820 bp      DNA      linear      PRI 27-MAR-2003
LOCUS      Homo sapiens 3 BAC RP11-513G11 (Roswell Park Cancer Institute Human
DEFINITION      BAC library) complete sequence.
ACCESSION      AC117469
VERSION      AC117469.3 GI:21240539
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 144820)
REFERENCE      1 (bases 1 to 144820)
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amarante,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbarta,T., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
            Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthett,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.U.,
            Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harrie,K., Hart,M., Havlak,P., Hawes,A., He,X.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Hollway,C.,
            Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Louissegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
            Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,N., Neal,D.,
            Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenko,S., Ogun,M., Okunodu,G., Otagunye,N.,
            Nickerson,E., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
            Pichens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rites,B.,
            Rojas,A., RojuboKan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S.,
            Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Swalek,A.,
            Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telford,B., Thomas,R., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Wellington,C., Williams,G.,
            Williamson,A., Wlaczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Zorrilla,S., Zorrilla,S.L., Weinstein,G. and
            Gibbs,K.
TITLE      Direct Submission
JOURNAL      Unpublished

```

```

REFERENCE      2 (bases 1 to 144820)
AUTHORS      Worley,K.C.
TITLE      Direct Submission
JOURNAL      Submitted (10-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE      3 (bases 1 to 144820)
AUTHORS      Worley,K.C.
TITLE      Direct Submission
JOURNAL      Submitted (26-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE      4 (bases 1 to 144820)
AUTHORS      Worley,K.C.
TITLE      Direct Submission
JOURNAL      Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE      5 (bases 1 to 144820)
AUTHORS      Worley,K.C.
TITLE      Direct Submission
JOURNAL      Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT      On May 29, 2002 this sequence version replaced gi:21206107.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

```

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

#### FEATURES

```

source
1..144820
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-513G11"
repeat_region
330..380
repeat_family="MIR"
repeat_region
572..903

```

repeat_region	/rpt_family="A1JMaS" complement(914. ,.993) /rpt_family="MER5A" 1025. ,.1204 /rpt_family="MER5A" complement(1213. ,.1374) /rpt_family="MER5S" complement(12387. ,.2682) /rpt_family="A1JubD" complement(31073. ,.3145) /rpt_family="MIR" complement(3664. ,.3791) /rpt_family="MIR" 3866. ,.3924 /rpt_family="TG(n)" 4339. ,.4364 /rpt_family="GGGA)n" complement(4656. ,.4766) /rpt_family="MIR" complement(4867. ,.4959) /rpt_family="L2" 4960. ,.5267 /rpt_family="A1JubD" complement(53268. ,.5423) /rpt_family="L2" 5454. ,.5554 /rpt_family="MER5A" 6001. ,.6220 /rpt_family="TCCA)n" 6794. ,.6818 /rpt_family="AT_rich" complement(8815. ,.8892) /rpt_family="L2" 9672. ,.9841 /rpt_family="MER5A" 11125. ,.11256 /rpt_family="L2" 12374. ,.12676 /rpt_family="A1JubD" complement(12768. ,.12877) /rpt_family="L2" 12896. ,.12447 /rpt_family="L2" complement(14248. ,.14540) /rpt_family="A1Juv" 14541. ,.14657 /rpt_family="L2" 14714. ,.14822 /rpt_family="MER97a" 14877. ,.15046 /rpt_family="L2" complement(15133. ,.15182) /rpt_family="L2" 15355. ,.15777 /rpt_family="L2" 17401. ,.17710 /rpt_family="A1Juv" 17728. ,.18039 /rpt_family="A1JusP" complement(18924. ,.19236) /rpt_family="A1JubO" 19377. ,.19672 /rpt_family="A1JusK" 20012. ,.20103 /rpt_family="CGGGG)n" 20434. ,.20465 /rpt_family="T)n" complement(21139. ,.21318) /rpt_family="A1JusK" complement(21319. ,.21624) /rpt_family="A1Juv" complement(21625. ,.21761) /rpt_family="A1JusK"
---------------	--

Query Match	Similarity	1.3%	Score	26;	DB	9;	Length	144820;
Beet local	Matches	26;	Conservative	0;	Mismatches	0;	Indels	0;
Db	2011	CCTACTAAATAATATAAATTAGCTG	2036					
		CCCTAAATAATATAAATTAGCTG	38424					
RESULT 73	AC021976	145086 bp	DNA	linear	HTG 07-JUL-2000			
LOCUS	Homo sapiens chromosome 17 clone RP11-578C11, WORKING DRAFT							
DEFINITION	SEQUENCE, 5 unordered pieces.							
ACCESSION	AC021976							
VERSION	AC021976.3	GI:7243941						
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
TITLE	Waterston,R.H.							
REFERENCE	The sequence of Homo sapiens clone							
JOURNAL	Unpublished							
AUTHORS	2 (bases 1 to 145086)							
TITLE	Waterston,R.H.							
JOURNAL	Direct Submission							
	Submitted (22-JAN-2000) Genome Sequencing Center, Washington							
	MO 63108, USA							
	On Mar 15, 2000 this sequence version replaced gi:7109655.							
COMMENT	----- Genome Center -----							
	Center: Washington University Genome Sequencing Center							
	Center code: WUGSC							
	Web site:http://genome.wustl.edu/gsc/index.shtml							
	----- Project Information -----							
	Center project name: H_NH0578C11							
	----- Summary Statistics -----							
	Sequencing vector: MJ3; 80%							
	Sequencing vector: plasmid; 20%							
	Chemistry: Dye-Primer ET; 80% of reads							
	Chemistry: Dye-terminator Big Dye; 20% of reads							
	Assembly program: Phrap; version 0.99019							
	Consensus quality: 141780 bases at least Q40							
	Consensus quality: 142468 bases at least Q30							
	Consensus quality: 142933 bases at least Q20							
	Insert size: 137000; agarose-fp							
	Insert size: 144686; sum-of-contigs							
	Quality coverage: 7.69 in Q20 bases; agarose-fp							
	Quality coverage: 7.34 in Q20 bases; sum-of-contigs							
	-----							
*	NOTE: This is a 'working draft' sequence. It currently							
*	consists of 5 contigs. The true order of the pieces							
*	is not known and their order in this sequence record is							
*	arbitrary. Gaps between the contigs are represented as							
*	'runs of N', but the exact sizes of the gaps are unknown.							
*	This record will be updated with the finished sequence							
*	as soon as it is available and the accession number will							
*	be preserved.							
*	1 5076: contig of 5076 bp in length							
*	5177 5176: gap of unknown length							
*	5177 5177: contig of 5012 bp in length							
*	14189 14288: gap of unknown length							
*	14289 23511: contig of 9223 bp in length							
*	23512 23611: gap of unknown length							
*	23612 63444: contig of 39833 bp in length							
*	63445 63544: gap of unknown length							
*	63545 145086: contig of 81542 bp in length.							
	Location/Qualifiers							
	1 .145086							

misc\_feature /organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/clone="RP11-578C11"  
1..5076  
/note="assembly\_name:Contig7  
clone\_end:SP6  
vector\_side:left"  
misc\_feature /note="assembly\_name:Contig8"  
5177..14188  
/note="assembly\_name:Contig9  
14289..23511  
/note="assembly\_name:Contig9  
clone\_end:T7  
vector\_side:left"  
misc\_feature 23612..63444  
/note="assembly\_name:Contig10"  
63545..145086  
misc\_feature /note="assembly\_name:Contig11"  
/note="assembly\_name:Contig11"  
ORIGIN  
Query Match 1.3%; Score 26; DB 2; Length 145086;  
Best Local Similarity 100.0%; Pred. No. 0.041;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2011 CCTACTAAATATATAATAGCTG 2036  
|||||  
Db 118938 CCTACTAAATATATAATAGCTG 118963  
|||||  
RESULT 74 150192 bp DNA linear PRI 11-AUG-2001  
AC008686/c Homo sapiens chromosome 19 clone CTB-5E10, complete sequence.  
LOCUS AC008686  
DEFINITION AC008686  
AC008686  
AC008686.8 GI:15148103  
VERSION HTG.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 150192)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submision  
JOURNAL Unpublished  
RECORD 2 (bases 1 to 150192)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submision  
JOURNAL Direct Submision  
SUBMITTED (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 150192)  
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submision  
SUBMITTED (11-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Aug 11, 2001 this sequence version replaced gi:15145649.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.hgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.1.  
FEATURES  
source 1..150192  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTB-5E10"  
ORIGIN  
Query Match 1.3%; Score 26; DB 9; Length 150192;  
Best Local Similarity 100.0%; Pred. No. 0.041;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2011 CCTACTAAATATATAATAGCTG 2036  
|||||  
Db 129265 CCTACTAAATATATAATAGCTG 129240  
|||||  
RESULT 75 151308 bp DNA linear PRI 01-MAY-2001  
AC015861  
LOCUS AC015861  
DEFINITION Homo sapiens, clone RP11-45A1, complete sequence.  
AC015861  
AC015861.7 GI:13376946  
VERSION HTG.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 151308)  
AUTHORS Birren, B., Linton, L., Nussbaum, C., and Lander, E.  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 151308)  
REFERENCE Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M.,  
Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Casle, A., Colangelo, M., Collins, S., Collymore, A.,  
Cooke, P., Deatellano, K., Dewar, K., Domino, M., Donlan, L., Doyle, M.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howard, J., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J.,  
Lehoczky, J., Lien, C., Locke, K., Macdonald, P., Margulis, N.,  
McEwan, P., McQuirk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamae, J.,  
Testaye, S., Tittel, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.  
TITLE Direct Submision  
JOURNAL Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 151308)  
REFERENCE Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Baethen, V., Boguslavsky, L., Boukhalter, B., Brown, A.,  
Camata, N., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,  
Collins, S., Collymore, A., Cooke, P., Deatellano, K., Dewar, K.,  
Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,  
Iliev, I., Johnson, R., Jones, C., Karatas, A., Labocque, K.,  
Lamazares, R., Landers, T., Lehoczky, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarthy, M.,  
McEwan, P., McKernan, K., McQuirk, R., Meldrum, J., Menus, L.,  
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Roman, J.,  
Retra, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Royan, J.,  
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapp, R., Seaman, S.,  
Severy, P., Sougnuez, C., Spencer, B., Strange-Thomann, N.,  
Stojanovic, N., Straus, N., Subramanian, A., Talamae, J., Testaye, S.,  
Theodore, J., Travers, M., Travis, N., Tittel, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
TITLE Direct Submision  
JOURNAL Submitted (01-MAY-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 18, 2001 this sequence version replaced gi:11693425.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIRB  
Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu

----- Project Information

Center project name: L616

Center clone name: 45\_A\_1

----- Location/Qualifiers

```

FEATURES
    source          1..151308
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="RP11-45A1"
                    /clone_11b="RPCT-11 Human Male BAC"
                    111..388
    repeat_region   /rpt_family="L1MA2"
                    1535..1569
                    /rpt_family="(TA)n"
    repeat_region   complement(1841..2183)
                    /rpt_family="THE1B"
                    2341..2407
    repeat_region   /rpt_family="AT_rich"
                    complement(2937..3120)
                    /rpt_family="L2"
    repeat_region   4384..4427
                    /rpt_family="AT_rich"
    repeat_region   6440..6491
                    /rpt_family="(CA)n"
    repeat_region   complement(6654..6963)
                    /rpt_family="L1ME1"
    repeat_region   6964..7009
                    /rpt_family="(TTTA)n"
    repeat_region   complement(7010..8089)
                    /rpt_family="L1ME1"
    repeat_region   9142..9168
                    /rpt_family="(CAA)n"
    repeat_region   9440..9589
                    /rpt_family="MLT1H1"
    repeat_region   9860..9967
                    /rpt_family="MLT1-int"
    repeat_region   10895..11276
                    /rpt_family="MLT1H1"
    repeat_region   12155..12781
                    /rpt_family="L2"
    repeat_region   complement(13046..14471)
                    /rpt_family="PTRS"
    repeat_region   14489..14523
                    /rpt_family="AT_rich"
    repeat_region   15102..15140
                    /rpt_family="AT_rich"
    repeat_region   17055..17507
                    /rpt_family="MER34"
    repeat_region   17684..17719
                    /rpt_family="(TG)n"
    repeat_region   18897..18988
                    /rpt_family="(TA)n"
    repeat_region   18991..19030
                    /rpt_family="AT_rich"
    repeat_region   21050..21109
                    /rpt_family="(CA)n"
    repeat_region   complement(21311..21881)
                    /rpt_family="L1MD3"
    repeat_region   complement(21927..22596)
                    /rpt_family="L1MCC"
    repeat_region   complement(22870..22950)
                    /rpt_family="L1MCC"
    repeat_region   22952..22982
                    /rpt_family="(TG)n"
    repeat_region   complement(23458..23855)
                    /rpt_family="MT1A1"
    repeat_region   complement(23908..24236)
                    /rpt_family="MLT1C"
    repeat_region   24237..24531
                    /rpt_family="AluY"
    repeat_region   complement(24532..24624)

```

```

    repeat_region   /rpt_family="MLT1C"
                    complement(24675..24848)
    repeat_region   /rpt_family="L1MCC"
                    26136..26159
    repeat_region   /rpt_family="AT_rich"
                    complement(26622..26758)
    repeat_region   /rpt_family="MIR3"
                    26968..27114
    repeat_region   /rpt_family="L1MCAa"
                    27175..27207
    repeat_region   /rpt_family="(TATATG)n"
                    27225..27426
    repeat_region   /rpt_family="L1MCAa"
                    complement(27911..27991)
    repeat_region   /rpt_family="L1MD3"
                    complement(27985..28166)
    repeat_region   /rpt_family="L1MCA"
                    complement(28196..28567)
    repeat_region   /rpt_family="L1MCAa"
                    complement(28835..29245)
    repeat_region   /rpt_family="MLT1H1"
                    29303..29609
    repeat_region   /rpt_family="L1ME2"
                    29622..29833
    repeat_region   /rpt_family="L1ME2"
                    29836..30115
    repeat_region   /rpt_family="L1MCI"
                    complement(30127..30964)
    repeat_region   /rpt_family="L1PA3"
                    30960..31796
    repeat_region   /rpt_family="L1PA3"
                    31785..32367
    repeat_region   /rpt_family="L1MCI"
                    32420..32901
    repeat_region   /rpt_family="L1MCI"
                    32902..32930
    repeat_region   /rpt_family="(TTA)n"
                    32931..32987
    repeat_region   /rpt_family="L1MCI"
                    32988..33288
    repeat_region   /rpt_family="AluY"
                    33289..34913
    repeat_region   /rpt_family="L1MCI"
                    34973..35504
    repeat_region   /rpt_family="L1ME2"
                    35505..35614
    repeat_region   /rpt_family="L2"
                    35695..35726
    repeat_region   /rpt_family="AT_rich"
                    35728..35840
    repeat_region   /rpt_family="L2"
                    complement(38093..38210)
    repeat_region   /rpt_family="L2"
                    38211..38509
    repeat_region   /rpt_family="AluSc"
                    complement(38510..39072)
    repeat_region   /rpt_family="L2"
                    complement(39208..39455)
    repeat_region   /rpt_family="HAL1b"
                    complement(39912..40034)
    repeat_region   /rpt_family="AluSg/x"
                    40894..40942
    repeat_region   /rpt_family="AT_rich"

```

Query Match 1.3% Score 26; DB 9; Length 151308;  
 Best Local Similarity 100.0%; Pred. No. 0.041;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2011 CCTACTAAATATATATATAGCTG 2036  
 DB 91697 CCTACTAAATATATATATAGCTG 91722

RESULT 76  
AC145879/C 151392 bp DNA linear PRI 29-OCT-2003

LOCUS AC145879  
DEFINITION Pan troglodytes BAC clone RP43-20K4 from 7, complete sequence.  
ACCESSION AC145879  
VERSION AC145879.2 GI:37951458  
KEYWORDS HTG;  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
REFERENCE AUTHORS  
TITLE The sequence of Pan troglodytes BAC clone RP43-20K4  
JOURNAL Unpublished (2001)  
AUTHORS Sulston, J.E. and Wilson, R.  
TITLE Sequencing of Pan troglodytes  
JOURNAL Unpublished (2001)  
AUTHORS Wilson, R.K.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA  
AUTHORS Wilson, R.K.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-2003) Genetics, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA  
AUTHORS Wilson, R.K.  
TITLE Direct Submission  
JOURNAL Submitted (24-OCT-2003) Genetics, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA  
AUTHORS Wilson, R.  
TITLE Direct Submission  
JOURNAL Submitted (29-OCT-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT On Oct 24, 2003 this sequence version replaced gi:33386809.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@watson.wustl.edu  
----- Summary Statistics  
Center project name: C\_PT020K04  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see  
http://genome.wustl.edu

SOURCE INFORMATION:  
The RPI-1-3 BAC library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Cint', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen

```

(http://www.resgen.com) or Pieter de Jong and co-workers at
http://www.becpac.chori.org.

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1..151392
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="7"
/map="7"
/clone="RP43-20K4"
/clone_1ib="RPCT-43"
99339..99714
/notes="Sequence derived from PCR product of project DNA."
150865..150944
/notes="Sequence derived from one plasmid subclone."

ORIGIN

Query Match 1.3%; Score 26; DB 9; Length 151392;
Beat Local Similarity 100.0%; Pred. No. 0.041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2011 CCTACTAAATAATATAAATACTGTG 2036
|||||
Db 11380 CCTACTAAATAATATAAATACTGTG 11355

RESULT 77
AC017030/c 153148 bp DNA linear HTG 07-JUL-2000
LOCUS AC017030 Homo sapiens chromosome 19 clone RP11-285H8, WORKING DRAFT
DEFINITION SEQUENCE, 26 unordered pieces.
AC017030
AC017030 AC017030.4 GI:7230956
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
Organism
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 153148)
Waterston,R.H.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 153148)
Waterston,R.H.
Waterston,R.H.
Direct Submission
Submitted (09-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Mar 13, 2000 this sequence version replaced gi:7022019.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0285H08
----- Summary Statistics -----
Sequencing vector: M13; 85%
Sequencing vector: plasmid; 15%
Chemistry: Dye-Primer ET; 85% of reads
Chemistry: Dye-terminator Big Dye; 15% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 141534 bases at least Q40
Consensus quality: 145178 bases at least Q30
Consensus quality: 146980 bases at least Q20
Insert size: 171000; agarose-fp
Insert size: 150648; sum-of-contigs
Quality coverage: 4.09 in Q20 bases; agarose-fp
Quality coverage: 4.70 in Q20 bases; sum-of-contigs
-----

```

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1063: contig of 1063 bp in length
1064
1163: gap of unknown length
1164
2272: contig of 1109 bp in length
2273
2372: gap of unknown length
2373
4014: contig of 1642 bp in length
4015
4114: gap of unknown length
4115
5231: contig of 1117 bp in length
5232
5331: gap of unknown length
5332
7936: contig of 2605 bp in length
7937
8037: gap of unknown length
8038
9838: contig of 1802 bp in length
9839
9939
11921: contig of 1983 bp in length
11922
12021: gap of unknown length
12022
14249: contig of 2228 bp in length
14250
14349: gap of unknown length
14350
16974: contig of 2625 bp in length
16975
17074: gap of unknown length
17075
19409: contig of 2335 bp in length
19410
19509: gap of unknown length
19510
23043: contig of 3534 bp in length
23044
23143: gap of unknown length
23144
27568: contig of 4425 bp in length
27569
31622: gap of unknown length
31623
31722: contig of 3954 bp in length
31723
35164: contig of 3442 bp in length
35165
35264: gap of unknown length
35265
39685: contig of 4421 bp in length
39686
39785: gap of unknown length
39786
44143: contig of 4358 bp in length
44144
44244: gap of unknown length
44245
48478: contig of 4235 bp in length
48479
48578: gap of unknown length
48579
56908: contig of 8330 bp in length
56909
57008: gap of unknown length
57009
65030: contig of 8022 bp in length
65031
65130: gap of unknown length
65131
76558: contig of 11428 bp in length
76559
76658: gap of unknown length
76659
84912: contig of 8254 bp in length
84913
85012: gap of unknown length
85013
95387: contig of 10375 bp in length
95388
95487: gap of unknown length
95488
107504: contig of 12017 bp in length
107505
107604: gap of unknown length
107605
119850: contig of 12246 bp in length
119851
119950: gap of unknown length
119951
131781: contig of 11831 bp in length
131782
153148: gap of unknown length
153149
153148: contig of 21267 bp in length.
Location/Qualifiers
1..153148
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-285H8"
1..1063
/note="assembly_name:Contig7"
misc_feature
1164..2272
/note="assembly_name:Contig8"
misc_feature
2373..4014
/note="assembly_name:Contig9"
4115..5231
misc_feature

/note="assembly_name:Contig11"
5332..7936
/note="assembly_name:Contig12"
8037..9838
/note="assembly_name:Contig13"
9939..11921
/note="assembly_name:Contig14
clone_end:r7
vector_side:right"
12022..14249
/note="assembly_name:Contig15"
14350..16974
/note="assembly_name:Contig16"
17075..19409
/note="assembly_name:Contig17"
19510..23043
/note="assembly_name:Contig18"
23144..27568
/note="assembly_name:Contig19"
27669..31622
/note="assembly_name:Contig20"
31723..35164
/note="assembly_name:Contig21"
35265..39685
/note="assembly_name:Contig22"
39786..44143
/note="assembly_name:Contig23"
44244..48478
/note="assembly_name:Contig24"
48579..56908
/note="assembly_name:Contig25"
57009..65030
/note="assembly_name:Contig26"
65131..76558
/note="assembly_name:Contig27"
76659..84912
/note="assembly_name:Contig28"
85013..95387
/note="assembly_name:Contig29"
95488..107504
/note="assembly_name:Contig30"
107605..119850
/note="assembly_name:Contig31"
119951..131781
/note="assembly_name:Contig32"
131882..153148
/note="assembly_name:Contig33"

ORIGIN
Query Match 13%; Score 26; DB 2; Length 153148;
Best Local Similarity 100.0%; Pred.No. 0.041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATTAATAGCTG 2036
Db 65316 CCTACTAAATATTAATAGCTG 65291

RESULT 78
AC011471 154702 bp DNA linear PRI 22-DEC-2000
LOCUS Homo sapiens chromosome 19 clone CTC-503J8, complete sequence.
DEFINITION AC011471
ACCESSION AC011471.6 GI:11968298
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 154702)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Unpublished
```

REFERENCE 2 (bases 1 to 154702)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 154702)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 On Dec 22, 2000 this sequence version replaced gi:7711508.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 99.9% of Sequence;  
 Estimated Total Number of Errors is 0.6.

FEATURES  
 source  
 1. 154702  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="CTC-50308"

ORIGIN  
 Query Match 1.3%; Score 26; DB 9; Length 154702;  
 Best Local Similarity 100.0%; Pred. No. 0.041;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATAGCTG 2036  
 |||||  
 DB 147715 CCTACTAAATATATAAATAGCTG 147740

RESULT 79  
 AC026254 154906 bp DNA linear PRI 29-DEC-2002  
 LOCUS Homo sapiens chromosome 17, clone RP11-173M1, complete sequence.  
 DEFINITION AC026254  
 AC026254  
 VERSION AC026254.9 GI:27413913  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 154906)  
 Homo sapiens chromosome 17, clone RP11-173M1  
 Unpublished  
 2 (bases 1 to 154906)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G., Campolongo, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collumore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hages, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPherson, R., Melrim, J., Menus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission

JOURNAL Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 154906)  
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamt, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, J., Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission

JOURNAL Submitted (05-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 154906)  
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamt, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, J., Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission

JOURNAL Submitted (23-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 5 (bases 1 to 154906)  
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamt, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, J., Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission

JOURNAL Submitted (29-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 29, 2002 this sequence version replaced gi:21930266.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center

COMMENT



Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center Project name: L6643  
 Center clone name: 173\_M\_1  
 -----

FEATURES  
 source  
 Location/Qualifiers  
 1..154906  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-173M1"  
 /clone\_1b="RPci-11 Human Male BAC"  
 1..152  
 repeat\_region  
 /rpt\_family="L1M4"  
 229..345  
 /rpt\_family="L1M4"  
 repeat\_region  
 complement(344..479)  
 /rpt\_family="L1ME3"  
 538..837  
 repeat\_region  
 /rpt\_family="AluUb"  
 1352..1438  
 repeat\_region  
 /rpt\_family="L1PA12"  
 complement(12560..2601)  
 /rpt\_family="MLT1L"  
 repeat\_region  
 complement(2669..2843)  
 /rpt\_family="MLT1L"  
 repeat\_region  
 complement(2987..3100)  
 /rpt\_family="MIR3"  
 3128..3534  
 repeat\_region  
 /rpt\_family="MLT2D"  
 3981..4092  
 repeat\_region  
 /rpt\_family="LTR67"  
 complement(4101..4444)  
 /rpt\_family="THE1C"  
 4698..4801  
 repeat\_region  
 /rpt\_family="MIR3"  
 complement(5207..5514)  
 /rpt\_family="AluSx"  
 5633..5672  
 repeat\_region  
 /rpt\_family="AT\_rich"  
 complement(5689..5829)  
 /rpt\_family="AluUb"  
 5971..6266  
 repeat\_region  
 /rpt\_family="AluUb"  
 6276..6583  
 repeat\_region  
 /rpt\_family="AluJo"  
 complement(6642..6801)  
 /rpt\_family="MIR"  
 7475..7599  
 repeat\_region  
 /rpt\_family="MIR3"  
 8052..8117  
 repeat\_region  
 /rpt\_family="AluU/FRAM"  
 8242..8265  
 repeat\_region  
 /rpt\_family="AT\_rich"  
 8462..8638  
 repeat\_region  
 /rpt\_family="MIR"  
 9086..9439  
 repeat\_region  
 /rpt\_family="AluUb"  
 complement(10444..10561)  
 /rpt\_family="MIR"  
 10825..10943  
 repeat\_region  
 /rpt\_family="L2"  
 11572..11610  
 repeat\_region  
 /rpt\_family="L2"  
 11796..12108  
 repeat\_region  
 /rpt\_family="AluSx"  
 complement(12791..12882)  
 /rpt\_family="MERSB"

repeat\_region 12888..13214  
 /rpt\_family="AluSg"  
 repeat\_region 13778..13936  
 /rpt\_family="L2"  
 repeat\_region complement(14841..14892)  
 /rpt\_family="MIR"  
 repeat\_region complement(15134..15274)  
 /rpt\_family="LTR67"  
 repeat\_region complement(15436..15646)  
 /rpt\_family="LTR67"  
 repeat\_region complement(16688..16854)  
 /rpt\_family="AluUb"  
 repeat\_region 16927..16998  
 /rpt\_family="MLT1L"  
 repeat\_region complement(17012..17300)  
 /rpt\_family="AluSg"  
 repeat\_region complement(17369..17691)  
 /rpt\_family="MER2"

Query Match 1.3% Score 26; DB 9; Length 154906;  
 Best Local Similarity 100.0%; Pred.No. 0.041;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
 DB 143016 CCTACTAAATATATAAATTAGCTG 143041

RESULT 80  
 AC141275/c 155216 bp DNA linear HTG 11-MAR-2003  
 LOCUS Homo sapiens chromosome 16 clone RP11-208020, WORKING DRAFT  
 DEFINITION  
 SEQUENCE.  
 AC141275  
 AC141275.1 GI:28913055  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Sequencing of Human Chromosome 16  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 155216)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: http://www.jgi.doe.gov  
 -----

Project Information  
 Project Name: 486593  
 Center Project Name: RPci-11\_208020  
 Center clone name: RPci-11\_208020  
 -----

Summary Statistics  
 Consensus quality: 155216 bases at least Q40  
 Consensus quality: 155216 bases at least Q30  
 Consensus quality: 155216 bases at least Q20  
 Estimated insert size: 175000; agarose-fp estimation  
 Estimated insert size: 155216; sum-of-ct-ctigs estimation  
 Quality coverage: 13.78 in Q20 bases; agarose-fp estimation  
 Quality coverage: 15.54 in Q20 bases; sum-of-ct-ctigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will.

\* be preserved.  
1 155216: config of 155216 bp in length.

FEATURES  
source  
1.155216

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-208020"  
/clone\_lib="RP11 human BAC library 11"

## ORIGIN

Query Match 1.3%; Score 26; DB 2; Length 155216;  
Best Local Similarity 100.0%; Pred.No. 0.041;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
|||||  
DB 116957 CCTACTAAATATATAAATTAGCTG 116932

RESULT 81  
AC141313 155216 bp DNA linear HTG 11-MAR-2003  
LOCUS Homo sapiens chromosome 16 clone RP11-91N11, WORKING DRAFT  
DEFINITION  
SEQUENCE.  
AC141313  
AC141313.1 GI:28913093  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 155216)  
DOE Joint Genome Institute.  
Unpublished  
Sequencing of Human Chromosome 16  
2 (bases 1 to 155216)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 441632  
Center clone name: RP11-91N11  
-----

Summary Statistics  
Consensus quality: 155215 bases at least Q40  
Consensus quality: 155216 bases at least Q30  
Consensus quality: 155216 bases at least Q20  
Estimated insert size: 175000; agarose-fp estimation  
Estimated insert size: 155216; sum-of-ctnigs estimation  
Quality coverage: 9.23 in Q20 bases; agarose-fp estimation  
Quality coverage: 10.41 in Q20 bases; sum-of-ctnigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

## FEATURES

source  
1.155216  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

/chromosome="16"  
/clone="RP11-91N11"  
/clone\_lib="RP11 human BAC library 11"

## ORIGIN

Query Match 1.3%; Score 26; DB 2; Length 155216;  
Best Local Similarity 100.0%; Pred.No. 0.041;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
|||||  
DB 116957 CCTACTAAATATATAAATTAGCTG 116932

RESULT 82  
AL591720/c 156491 bp DNA linear PRI 16-NOV-2001  
LOCUS Human DNA sequence from clone RP11-28K12 on chromosome 1, complete  
DEFINITION  
sequence.  
AL591720  
AL591720.12 GI:16973945  
HTG.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Glithero, R.  
Direct Submission  
Submitted (16-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk  
On Nov 17, 2001 this sequence version replaced gi:15723811.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr1>  
RP11-28K12 is from the library RP11-11.1 constructed by the group  
of Plier de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone  
RP11-28K12. It may be shorter because we sequence overlapping  
sections only once, except for a short overlap.  
The true left end of clone RP11-28K12 is at 1 in this sequence. The  
true left end of clone RP11-205P11 is at 127031 in this sequence.  
The true left end of clone RP4-654H19 is at 154492 in this  
sequence.

## FEATURES

source  
1.156491  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-28K12"

misc\_feature /clone\_11b="RPCT-11.1"  
 68205..68361  
 /note="Sequence from overlapping clone Rp11-800ES  
 (AL603915). Assembly confirmed by restriction digest."  
 misc\_feature 87845..87850  
 /note="Sequence from overlapping clone Rp11-800ES  
 (AL603915). Assembly confirmed by restriction digest."

Query Match 1.3%; Score 26; DB 9; Length 156491;  
 Best Local Similarity 100.0%; Pred. No. 0.041;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2011 CCTACTAAATATATAATAGCTG 2036  
 DB 129224 CCTACTAAATATATAATAGCTG 129199

RESULT 83  
 AC009335/c AC009335 159611 bp DNA linear PRI 20-FEB-2002  
 LOCUS Homo sapiens chromosome 17, clone RP11-2002, complete sequence.  
 AC009335  
 AC009335.10 GI:18767554  
 VERSION HTG.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 JOURNAL Unpublished  
 TITLE 2 (bases 1 to 159611)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,  
 Casale, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., DeRellano, K., Depayre, E., Devon, K., Dewar, K.,  
 Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,  
 Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,  
 Hagos, B., Heath, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,  
 Karas, A., Lebecky, J., Lieu, C., Locke, K., MacDonald, P.,  
 Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,  
 Melgrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,  
 Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,  
 Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,  
 Testafave, S., Tornella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,  
 Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

TITLE Direct Submission  
 JOURNAL Submitted (16-AUG-1999) Whitehead Institute/MIT Center for Genome  
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA  
 AUTHORS 3 (bases 1 to 159611)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukgealter, B.,  
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,  
 Chopel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
 Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karas, A., Kellis, C., Lacombe, K., Lamazares, R.,  
 Landers, T., Lebecky, J., Levine, R., Liu, G., Maclean, C.,  
 Macdonald, P., McKernan, K., Melgrim, J., Menus, L., Mihova, T.,  
 Mleang, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Phukhang, P., Pierre, N., Pollara, V., Raymond, C.,  
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,  
 Rossetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,  
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Testafave, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

TITLE  
 JOURNAL  
 COMMENT  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (20-FEB-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Feb 20, 2002 this sequence version replaced gi:17975366.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seg.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Center project name: L927  
 Center clone name: 20\_02

FEATURES  
 source Location/Qualifiers  
 1..159611  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-2002"  
 /clone\_11b="RPCT-11 Human Male BAC"  
 281..357  
 /rpt\_family="MIR"  
 441..581  
 /rpt\_family="MIR"  
 681..712  
 /rpt\_family="(A)n"  
 986..1024  
 /rpt\_family="L2"  
 1025..1253  
 /rpt\_family="AluJo"  
 1969..2291  
 /rpt\_family="AluSg"  
 complement(2879..3176)  
 /rpt\_family="AluY"  
 3487..3798  
 /rpt\_family="AluSg"  
 complement(4041..4262)  
 /rpt\_family="L1M4"  
 complement(4263..4571)  
 /rpt\_family="AluY"  
 complement(4572..4892)  
 /rpt\_family="L1M4"  
 complement(4893..5228)  
 /rpt\_family="AluY"  
 complement(5229..5849)  
 /rpt\_family="L1M4"  
 6189..6436  
 /rpt\_family="L1ME1"  
 6435..6744  
 /rpt\_family="AluSc"  
 6892..6917  
 /rpt\_family="(TG)n"  
 7158..7662  
 /rpt\_family="L1MC/D"  
 7668..7976  
 /rpt\_family="AluSg"  
 8015..8093  
 /rpt\_family="MERL16"  
 8094..8250  
 /rpt\_family="L1MC/D"  
 8365..8387  
 /rpt\_family="(TA)n"  
 8388..8545  
 /rpt\_family="MER47A"  
 complement(8546..8830)  
 /rpt\_family="AluSg"

```

repeat_region      8831..9048
                    /rpt_family="MER47A"
repeat_region      9084..9144
                    /rpt_family="Tigser3(Golem)"
repeat_region      /rpt_family="Tigser3(Golem)"
                    complement(9179..9482)
repeat_region      9483..9755
                    /rpt_family="AluSg"
repeat_region      9908..10232
                    /rpt_family="Tigser3(Golem)"
                    /rpt_family="AluSx"
repeat_region      10285..10600
                    /rpt_family="AluSg"
repeat_region      10763..10971
                    /rpt_family="LTR33"
repeat_region      11106..11435
                    /rpt_family="MER39"
repeat_region      /rpt_family="MER39"
                    complement(11430..11882)
repeat_region      /rpt_family="MER39"
                    complement(11430..11882)
repeat_region      11883..12186
                    /rpt_family="AluSg"
                    complement(12187..12343)
repeat_region      /rpt_family="MER39"
                    complement(14147..14278)
repeat_region      14399..14695
                    /rpt_family="MIR"
repeat_region      15074..15276
                    /rpt_family="AluSg"
                    /rpt_family="MIR"
repeat_region      /rpt_family="MIR"
                    complement(15283..15586)
repeat_region      /rpt_family="AluSg"
                    complement(15638..15941)
repeat_region      /rpt_family="AluSg"
                    complement(15934..15938)
                    /note="<30 qual SNGL region"
repeat_region      16351..16654
                    /rpt_family="AluSg"
                    complement(16724..16925)
repeat_region      /rpt_family="LTR67"
                    complement(17081..17371)
repeat_region      /rpt_family="AluSg"
                    /rpt_family="L2"
                    complement(17835..18149)
repeat_region      18182..18262
                    /rpt_family="MER72"
                    /rpt_family="MER72"
repeat_region      18623..19175
                    /rpt_family="MER72"
                    /rpt_family="MER72"
repeat_region      19176..19476
                    /rpt_family="AluSx"
                    /rpt_family="MER72"
repeat_region      19477..19529
                    /rpt_family="MER72"
                    /rpt_family="MER21B"
repeat_region      19650..20360
                    /rpt_family="MER21B"
                    /rpt_family="MER21B"
repeat_region      20926..21230
                    /rpt_family="AluY"
                    /rpt_family="AluY"
repeat_region      21381..21569
                    /rpt_family="HAL1"
                    complement(21570..21758)
repeat_region      /rpt_family="LTPA6"
                    /rpt_family="LTPA6"
repeat_region      21759..21935
                    /rpt_family="HAL1"
                    /rpt_family="HAL1"
repeat_region      21925..22437
                    /rpt_family="HAL1"
                    /rpt_family="HAL1"
repeat_region      22606..22887
                    /rpt_family="AluY"
                    /rpt_family="AluY"
repeat_region      23113..23427
                    /rpt_family="AluSg"
                    /rpt_family="AluSg"
repeat_region      23464..23490
                    /rpt_family="AT_rich"
repeat_region      23493..23801

```

```

Query Match      1.3%; Score 26; DB 9; Length 159611;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2011 CCTACTAAATATATAAATTAGCTG 2036
Db      52619 CCTACTAAATATATAAATTAGCTG 52594

RESULT 84
AC009781/c
LOCUS
DEFINITION
AC009781
VERSION
AC009781.9 GI:20127897
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbata,J., Benton,J., Blinage,K., Blankenburg,K., Bonin,D.,
Bouck,J., Bowie,S., Brieleva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.P., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,T., Foster,P., Frantz,P.,
Gabel,S., Gao,J., Garcia,A., Garner,T., Garza,N., Gall,R.,
Gibbs,J., Givara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Hayak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hughes,M., Holloway,C., Hollins,B.,
Homel,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.V., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,B.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojubenkan,I., Rolfe,M., Ruhl,S., Savary,G.,
Scherrer,S., Scott,G., Shen,H., Shoohtlari,N., Sisson,I.,
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tameisa,K., Tameisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telirod,B., Thomas,N., Thomas,S.,
Umanli,K., Vaequez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 159693)
REFERENCE
Worley,K.C.
Direct Submission
Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 159693)

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE

```

## AUTHORS

Worley, K.C.  
 Direct Submission  
 Submitted (21-SEP-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Apr 10, 2002 this sequence version replaced gi:9966736.

## COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Drafting Center Code: BCM

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HMLS

Center clone name: RP11-256J12

----- Summary Statistics

Sequencing vector: M13;

Chemistry: Dye-terminator Big Dye; 87% of reads

Assembly program: Phrap; version 0.990326

Consensus quality: 145966 bases at least Q40

Consensus quality: 151155 bases at least Q30

Consensus quality: 154120 bases at least Q20

Estimated insert size: 166126; agarose-fp estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 13 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

-----

1 2937: contig of 2937 bp in length

2938 3037: gap of unknown length

3038 7655: contig of 4618 bp in length

7656 7755: gap of unknown length

7756 14135: contig of 6380 bp in length

14136 14235: gap of unknown length

14236 23766: contig of 9531 bp in length

23767 23866: gap of unknown length

23867 35696: contig of 11830 bp in length

35697 35796: gap of unknown length

35797 45824: contig of 10028 bp in length

45825 45924: gap of unknown length

45925 59028: contig of 13104 bp in length

59029 59128: gap of unknown length

59129 72237: contig of 13109 bp in length

72238 72337: gap of unknown length

72338 85979: contig of 13642 bp in length

85980 86079: gap of unknown length

86080 99513: contig of 13434 bp in length

99514 99613: gap of unknown length

99614 119656: contig of 20043 bp in length

119657 119756: gap of unknown length

119757 139833: contig of 20077 bp in length

139834 139933: gap of unknown length

139934 159693: contig of 19760 bp in length.

## FEATURES

## SOURCE

1. 159693  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-256J12"

## ORIGIN

Query Match 1.3%; Score 26; DB 2; Length 159693;  
 Best Local Similarity 100.0%; Pred. No. 0.041;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

2011 CCTACTAAATATTAATTCGCTG 2036

## Db

49166 CCTACTAAATATTAATTCGCTG 49141

## RESULT 85

## AC022248

## LOCUS

## DEFINITION

## AC022248

## VERSION

## AC022248.2

## KEYWORDS

## HTG; HTGS\_PHASE1; HTGS\_DRAFT.

## SOURCE

## Homo sapiens

## ORGANISM

## Homo sapiens

## REFERENCE

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,

## Birken, B., Linton, L., Nusbaum, C., Lander, E., Adrahan, H., Allen, N.,



JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 163652)  
 AUTHORS Green, E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JUL-2003) NIH Intramural Sequencing Center, 8717  
 GroveMont Circle, Gaithersburg, MD 20877, USA  
 REFERENCE 3 (bases 1 to 163652)  
 AUTHORS Green, E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-OCT-2003) NIH Intramural Sequencing Center, 8717  
 GroveMont Circle, Gaithersburg, MD 20877, USA  
 COMMENT On Oct 17, 2003 this sequence version replaced gi:3354320.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: http://www.nisc.nih.gov  
 Contact: nisc.zoo@nhgri.nih.gov  
 ----- Project Information  
 Center project name: eex  
 Center clone name: 030F21

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones; overlaps with neighboring clones; alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies. The low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

## ----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 160056 bases at least Q40  
 Consensus quality: 160943 bases at least Q30  
 Consensus quality: 161649 bases at least Q20  
 Insert size: 148000; agarose-fp  
 Insert size: 162052; sum-of-contigs  
 Quality coverage: 9.3ix in Q20 bases; agarose-fp  
 Quality coverage: 8.50x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 17 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

```

* 1 9579: contig of 9579 bp in length
* 9580 9679: gap of unknown length
* 9680 15619: contig of 5940 bp in length
* 15620 15719: gap of unknown length
* 15720 18172: contig of 2453 bp in length
* 18173 18272: gap of unknown length
* 18273 27917: contig of 9645 bp in length
* 27918 28017: gap of unknown length
* 28018 30392: contig of 2375 bp in length
* 30393 30492: gap of unknown length
* 30493 33736: contig of 3244 bp in length
* 33737 33836: gap of unknown length
* 33837 36525: contig of 2689 bp in length
* 36526 36625: gap of unknown length
* 36626 47269: contig of 10644 bp in length
* 47270 47369: gap of unknown length
* 47370 54198: contig of 6829 bp in length
* 54199 54298: gap of unknown length
* 54299 54299: contig of 3277 bp in length

```

```

* 57576 57675: gap of unknown length
* 57676 82147: contig of 24472 bp in length
* 82148 82247: gap of unknown length
* 82248 96812: contig of 14565 bp in length
* 96813 96912: gap of unknown length
* 96913 138261: contig of 41349 bp in length
* 138262 138361: gap of unknown length
* 138362 153193: contig of 14632 bp in length
* 153194 153293: gap of unknown length
* 153294 160733: contig of 7440 bp in length
* 160734 160833: gap of unknown length
* 160834 162817: contig of 1984 bp in length
* 162818 162918: gap of unknown length
* 162918 163652: contig of 735 bp in length.
Location/Qualifiers
1. 163652

```

## FEATURES

```

source
1. 163652
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-30F21"
/clone_1fb="RP43"
1. 29185
/note="clone overlaps with GenBank Accession Number
AC096679 clone RP43-14213 (center project name cen)"
1. 9579
/note="assembly_fragment
clone_end:17
vector_side:left"
misc_feature
/note="assembly_fragment"
15720..18172
/note="assembly_fragment"
18273..27917
/note="assembly_fragment"
28018..30392
/note="assembly_fragment"
30493..33736
/note="assembly_fragment"
33837..36525
/note="assembly_fragment"
36626..47269
/note="assembly_fragment"
47370..54198
/note="assembly_fragment"
54299..54299
/note="assembly_fragment"
54299..57575
/note="assembly_fragment"
57676..82147
/note="assembly_fragment"
82248..96812
/note="assembly_fragment"
96913..138261
/note="assembly_fragment"
138362..153193
/note="assembly_fragment"
153294..160733
/note="assembly_fragment"
160834..162817
/note="assembly_fragment"
162918..163652
/note="assembly_fragment
clone_end:SP6
vector_side:right"

```

## ORIGIN

```

Query Match 1.3%; Score 26; DB 2; Length 163652;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2011 CCTACTAAATATATAAATTAGCTG 2036
DB 149598 CCTACTAAATATATAAATTAGCTG 149573

```

```

RESULT 88
AC016034          165106 bp   DNA       1linear   HTG 12-MAR-2000
LOCUS
DEFINITION
Homo sapiens clone RP11-1516, WORKING DRAFT SEQUENCE, 20 unordered
pieces.
AC016034
AC016034.2   GI:7230105
VERSION
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 165106)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-1516
Unpublished
2 (bases 1 to 165106)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Donno,M., Donegan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galsgan,J., Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lohoczky,J., Liu,C., Locke,K., Macdonald,P., Margulis,N.,
McEwan,P., McGuirk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tjirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wymann,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:6453386.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3472
Center clone name: 15 I 6
----- Summary Statistics
Sequencing vector: M13; M77815, 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 119639 bases at least Q40
Consensus quality: 142460 bases at least Q30
Consensus quality: 155240 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 163206; sum-of-ctrls
Quality coverage: 4.2 in Q20 bases; agarose-fp
Quality coverage: 4.5 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1259: contig of 1259 bp in length
* 1260 1359: gap of 100 bp
* 3153 3153: contig of 1794 bp in length
* 3253 3253: gap of 100 bp
* 4998 4998: contig of 1745 bp in length
* 5098 5098: gap of 100 bp
* 7142 7142: contig of 2044 bp in length

```

```

* 7143 7242: gap of 100 bp
* 7243 10682: contig of 3440 bp in length
* 10683 10782: gap of 100 bp
* 10783 14440: contig of 3658 bp in length
* 14441 14540: gap of 100 bp
* 14541 18349: contig of 3809 bp in length
* 18350 18449: gap of 100 bp
* 18450 23099: contig of 4650 bp in length
* 23100 23199: gap of 100 bp
* 23200 26844: contig of 3645 bp in length
* 26845 26944: gap of 100 bp
* 26945 30699: contig of 3755 bp in length
* 30700 30799: gap of 100 bp
* 30800 35612: contig of 4813 bp in length
* 35613 35712: gap of 100 bp
* 35713 40942: contig of 5230 bp in length
* 40943 41042: gap of 100 bp
* 41043 48694: contig of 7652 bp in length
* 48695 48794: gap of 100 bp
* 48795 56273: contig of 7479 bp in length
* 56274 56374: gap of 100 bp
* 56374 63898: contig of 7525 bp in length
* 63899 63998: gap of 100 bp
* 63999 73416: contig of 9418 bp in length
* 73417 73516: gap of 100 bp
* 73517 84930: contig of 11414 bp in length
* 84931 85030: gap of 100 bp
* 85031 100550: contig of 1520 bp in length
* 100551 100650: gap of 100 bp
* 100651 128624: contig of 27974 bp in length
* 128625 128724: gap of 100 bp
* 128725 165106: contig of 36382 bp in length.

```

## FEATURES

## source

```

misc_feature
1.1259
/note="assembly_fragment"
1360..3153
/note="assembly_fragment"
3254..4998
/note="assembly_fragment"
5099..7142
/note="assembly_fragment"
7243..10682
/note="assembly_fragment"
10783..14440
/note="assembly_fragment"
14541..18349
/note="assembly_fragment"
18450..23099
/note="assembly_fragment"
23200..26844
/note="assembly_fragment"
clone_end:r7
vector_side:right"
26945..30699
/note="assembly_fragment"
30800..35612
/note="assembly_fragment"
35713..40942
/note="assembly_fragment"
clone_end:sp6
vector_side:right"
41043..48694
/note="assembly_fragment"
48795..56273
/note="assembly_fragment"
56374..63898
/note="assembly_fragment"

```



```

misc_feature      63999..73416
                    /note="assembly_fragment"
misc_feature      73517..84930
                    /note="assembly_fragment"
misc_feature      85031..100550
                    /note="assembly_fragment"
misc_feature      100651..128624
                    /note="assembly_fragment"
misc_feature      128725..165106
                    /note="assembly_fragment"

ORIGIN
Query Match      1.3% Score 26; DB 2; Length 165106;
Best Local Similarity 100.0%; Pred.No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      2011 CCTACTAAAAATATTAATAGCTG 2036
        |||||
Db      24310 CCTACTAAAAATATTAATAGCTG 24335

RESULT 89
LOCUS      AC027575             165447 bp    DNA    linear    PRI 16-AUG-2002
DEFINITION Homo sapiens chromosome 18, clone RP11-162A12, complete sequence.
ACCESSION      AC027575
VERSION      AC027575.13    GI:22267725
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 165447)
AUTHORS      Birren,B., Nubbaum,C., Lander,E., Abraham,H., Allen,N.,
               Bogsuslavsky,L., Boughalter,B., Brown,A., Burkett,G.,
               Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
               Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
               Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitchugh,W., Gage,D.,
               Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
               Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
               Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
               Klein,J., Laroque,K., Lamazares,R., Lander,E., Lehocsky,J.,
               Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
               McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McNeer,R.,
               Melrim,J., Menue,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
               Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
               O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
               Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
               Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
               Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
               Testafaye,S., Theodore,J., Titrill,A., Travers,M., Trigilio,J.,
               Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
               Young,G., Zainoun,J., Zimmer,A. and Zody,M.
               Direct Submission
               Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
               3 (bases 1 to 165447)
REFERENCE      3 (bases 1 to 165447)
AUTHORS      Birren,B., Nubbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
               Barina,N., Bastien,V., Bloom,T., Bogsuslavsky,L., Boughalter,B.,
               Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
               Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
               Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
               Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
               Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
               Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
               McCarthy,M., Melrim,J., Menue,L., Mihova,T., Miranda,C., Mlenga,V.,
               Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
               O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
               Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
               Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
               Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
               Testafaye,S., Theodore,J., Titrill,A., Travers,M., Trigilio,J.,
               Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
               Zimmer,A. and Zody,M.
               Direct Submission
               Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
               4 (bases 1 to 165447)
REFERENCE      4 (bases 1 to 165447)
AUTHORS      Birren,B., Nubbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
               Barina,N., Bastien,V., Bloom,T., Bogsuslavsky,L., Boughalter,B.,
               Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
               Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
               Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
               Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
               Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
               Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
               McCarthy,M., Melrim,J., Menue,L., Mihova,T., Miranda,C., Mlenga,V.,
               Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
               O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
               Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
               Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
               Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
               Testafaye,S., Theodore,J., Titrill,A., Travers,M., Trigilio,J.,
               Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
               Zimmer,A. and Zody,M.
               Direct Submission
               Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
               On Aug 16, 2002 this sequence version replaced gi:2213714.
               All repeats were identified using RepeatMasker:
               Smit, A.F.A. & Green, P. (1996-1997)
               http://ftp.genome.washington.edu/RM/RepeatMasker.html
               ----- Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence submissions@genome.wi.mit.edu
               ----- Project Information
               Center project name: 162_A_12
               Center clone name: 162_A_12
               -----
               Location/Qualifiers
               1..165447
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="18"
               /map="18"
               /clone="RP11-162A12"
               /clone_lib="RPCT-11 Human Male BAC"
               /complement(1..399)
               /rpt_family="MER65-int"
               1077..1371
               /rpt_family="AluX"
               /complement(1545..1979)
               /rpt_family="MER65A"
               /complement(2488..2806)
               /rpt_family="AluY"
               2518..2527
               /note="<30 qual SNGL region"
               2580..2584
               /note="<30 qual SNGL region"
               2692..2698
               /note="<30 qual SNGL region"
               2719..2723
               /note="<30 qual SNGL region"
               /complement(2850..3152)
               /rpt_family="AluYo"
               3162..3183

```

```

O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 165447)
REFERENCE      4 (bases 1 to 165447)
AUTHORS      Birren,B., Nubbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
               Barina,N., Bastien,V., Bloom,T., Bogsuslavsky,L., Boughalter,B.,
               Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
               Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
               Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
               Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
               Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
               Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
               McCarthy,M., Melrim,J., Menue,L., Mihova,T., Miranda,C., Mlenga,V.,
               Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
               O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
               Pisanic,C., Pollara,V., Raymond,C., Retta,R., Rise,C., Rogov,P.,
               Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
               Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
               Testafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
               Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
               Zembek,L., Zimmer,A. and Zody,M.
               Direct Submission
               Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
               On Aug 16, 2002 this sequence version replaced gi:2213714.
               All repeats were identified using RepeatMasker:
               Smit, A.F.A. & Green, P. (1996-1997)
               http://ftp.genome.washington.edu/RM/RepeatMasker.html
               ----- Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence submissions@genome.wi.mit.edu
               ----- Project Information
               Center project name: 162_A_12
               Center clone name: 162_A_12
               -----
               Location/Qualifiers
               1..165447
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="18"
               /map="18"
               /clone="RP11-162A12"
               /clone_lib="RPCT-11 Human Male BAC"
               /complement(1..399)
               /rpt_family="MER65-int"
               1077..1371
               /rpt_family="AluX"
               /complement(1545..1979)
               /rpt_family="MER65A"
               /complement(2488..2806)
               /rpt_family="AluY"
               2518..2527
               /note="<30 qual SNGL region"
               2580..2584
               /note="<30 qual SNGL region"
               2692..2698
               /note="<30 qual SNGL region"
               2719..2723
               /note="<30 qual SNGL region"
               /complement(2850..3152)
               /rpt_family="AluYo"
               3162..3183

```

```

/rpt_family="AT_rich"
repeat_region 4070. .4092
/rpt_family="AT_rich"
repeat_region complement(4278.4579)
/rpt_family="AlusX"
repeat_region 4758. .5272
/rpt_family="L1MC1"
repeat_region 5273. .5623
/rpt_family="Zaphod"
repeat_region complement(7489.7646)
/rpt_family="L1MB3"
repeat_region complement(9259. .9506)
/rpt_family="L1PA5"
repeat_region complement(9532. .9839)
/rpt_family="AluY"
repeat_region complement(9840. .9905)
/rpt_family="L1P2"
repeat_region complement(9906. .10208)
/rpt_family="AluY"
repeat_region complement(10209. .10580)
/rpt_family="L1P2"
repeat_region complement(11831. .12122)
/rpt_family="HAL1"
repeat_region complement(12123. .12445)
/rpt_family="AluYb"
repeat_region complement(12446. .12738)
/rpt_family="HAL1"
repeat_region complement(12739. .13007)
/rpt_family="AlusX"
repeat_region complement(13008. .13373)
/rpt_family="HAL1"
repeat_region complement(13432. .13660)
/rpt_family="MERS8A"
repeat_region complement(14064. .14362)
/rpt_family="AlusX"
repeat_region complement(14366. .14481)
/rpt_family="HAL1"
repeat_region 15234. .15277
/rpt_family="CT-rich"
repeat_region 15278. .15305
/rpt_family="CA)n"
repeat_region complement(15671. .15826)
/rpt_family="MIR"
repeat_region 15996. .16140
/rpt_family="MIR3"
repeat_region 16330. .16671
/rpt_family="L2"
repeat_region complement(16757. .18359)
/rpt_family="L1MB5"
repeat_region 18448. .18661
/rpt_family="L2"
repeat_region complement(18969. .19033)
/rpt_family="MLT1A1"
repeat_region complement(19034. .19329)
/rpt_family="AluO"
repeat_region complement(19330. .19421)
/rpt_family="MLT1A1"
repeat_region 21073. .21572
/rpt_family="L1MD2"
repeat_region 21574. .21633
/rpt_family="MLT1J"
repeat_region 21931. .22090
/rpt_family="TAGTG)n"
repeat_region 22195. .22494
/rpt_family="AluY"
repeat_region complement(22504. .22709)
/rpt_family="L1MB3"
repeat_region complement(22742. .22784)
/rpt_family="MLT1B"
repeat_region 22785. .22936
/rpt_family="AlusG/x"
repeat_region 22937. .23132
/rpt_family="AlusG/x"

```

---

```

repeat_region complement(23133. .23515)
repeat_region /rpt_family="MLT1B"
/rpt_family="L1MD2"

Query Match 1.3%; Score 26; DB 9; Length 165447;
Best Local Similarity 100.0%; Pred.No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTTAAATAATTAATTTAGCTG 2036
DB 118944 CCTACTTAAATAATTAATTTAGCTG 118969

RESULT 90
AC023301/c 167190 bp DNA linear PRI 23-JUN-2002
LOCUS Homo sapiens chromosome 18, clone RP11-713C5, complete sequence.
DEFINITION AC023301
ACCESSION AC023301.5 GI:18266658
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 167190)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 18, clone RP11-713C5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 167190)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beda, F., Boguslavsky, L.,
Bouckhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeRellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,
Fenster, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,
Galgan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Landers, T., Lagocque, K., Lehoczy, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
McEwan, P., McGuirk, A., McKernan, K., McPheters, R., Melidim, J.,
Menas, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M.,
Peterson, K., Pierre, N., Pisani, C., Pollara, A., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirell, A.,
Travers, M., Trigglio, J., Vassiliev, H., Vei, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and
Zody, M.
DIRECT SUBMISSION
JOURNAL Submitted (11-FEB-2000) Whitehead Institute/MIT Center for Genome
RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 167190)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A.,
Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galgan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kelis, C., LaRoque, K.,
Lamasz, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Melidim, J.,
Menas, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
Notou, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, A.,
Raymond, C., Recta, R., Rieback, M., Riley, R., Riese, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

```

STRAUS, N., SUBRAMANIAN, A., TALAMAS, J., TRESFAYE, S., THEODORE, J.,  
TOPHAM, K., TRAVERS, M., TRAVIS, N., TRIGLILO, J., VASSILIEV, H.,  
VIEL, R., VO, A., WILSON, B., WU, X., WYMAN, D., YE, W.-J., YOUNG, G.,  
ZAHNOUN, J., ZEMBKE, L., ZIMMER, A. and ZODY, M.  
Direct Submission  
Submitted (23-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 23, 2002 this sequence version replaced gi:17432878.

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L4266
Center clone name: 713_C_5
-----
FEATURES
Location/Qualifiers
1..167130
source

```

repeat_region	complement(135696. .16017)
repeat_region	/rpt_family="MLT1G1"
repeat_region	17000. .17045
repeat_region	/rpt_family="CA-rich"
repeat_region	complement(18037. .18129)
repeat_region	/rpt_family="L1MEC"
repeat_region	18165. .18165
repeat_region	/rpt_family="AT-rich"
repeat_region	complement(18184. .18861)
repeat_region	/rpt_family="L1MEC"
repeat_region	complement(19203. .19788)
repeat_region	/rpt_family="L1MEC"
repeat_region	20823. .20956
repeat_region	/rpt_family="MIR"
repeat_region	complement(21276. .21575)
repeat_region	/rpt_family="AluO"
repeat_region	complement(23802. .23963)
repeat_region	/rpt_family="L1M4"
repeat_region	24679. .24677
repeat_region	/rpt_family="CA)n"
repeat_region	24700. .24706
repeat_region	/note="<30 qual SNGL region"
repeat_region	25240. .25284
repeat_region	/rpt_family="CA)n"
repeat_region	25342. .25381
repeat_region	/rpt_family="TG)n"
repeat_region	27423. .27650
repeat_region	/rpt_family="MIR"
repeat_region	complement(28462. .28534)
repeat_region	/rpt_family="MIR3"
repeat_region	complement(28829. .29260)
repeat_region	/rpt_family="Char1el"
repeat_region	complement(29261. .29556)
repeat_region	/rpt_family="AluSp"
repeat_region	complement(29557. .31737)
repeat_region	/rpt_family="Char1el"
repeat_region	complement(32019. .32154)
repeat_region	/rpt_family="MLT1J2"
repeat_region	complement(32314. .33382)
repeat_region	/rpt_family="L2"
repeat_region	complement(33289. .34978)
repeat_region	/rpt_family="Char1el"
repeat_region	complement(35008. .35522)
repeat_region	/rpt_family="L2"
repeat_region	36317. .36317
repeat_region	/rpt_family="AluYa5"
repeat_region	37201. .37498
repeat_region	/rpt_family="L1MA8"
repeat_region	complement(39392. .39692)
repeat_region	/rpt_family="MER102b"
repeat_region	complement(40351. .40612)
repeat_region	/rpt_family="L2"
repeat_region	complement(40649. .40945)
repeat_region	/rpt_family="AluX"
repeat_region	42186. .42251
repeat_region	/rpt_family="L2"
repeat_region	42302. .42371
repeat_region	/rpt_family="L3"
repeat_region	complement(42607. .42691)
repeat_region	/rpt_family="MIR"
repeat_region	42725. .42770
repeat_region	/rpt_family="AT-rich"
repeat_region	complement(42758. .42877)
repeat_region	/rpt_family="L1MA7"
repeat_region	43046. .43067
repeat_region	/rpt_family="AT-rich"
repeat_region	43133. .43556
repeat_region	/rpt_family="MLT1E1"
repeat_region	43847. .43895
repeat_region	/rpt_family="(TG)n"

Query Match	1.3%;	Score 26;	DB 9;	Length 167190;
Best Local Similarity	100.0%;	Pred. No. 0.04;		

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATAATTAATAATAGCTG 2036  
 |||  
 Db 74429 CCTACTAAATAATTAATAATAGCTG 74404

RESULT 91  
 AC132072/c

LOCUS AC132072 169904 bp DNA linear HTG 11-DEC-2002  
 DEFINITION Papio anubis clone RP41-162P7, WORKING DRAFT SEQUENCE, 9 unordered  
 pieces.

ACCESSION AC132072 GI:26449068  
 VERSION HTG: HTGS PHASE1: HTGS DRAFT.  
 KEYWORDS Papio anubis (olive baboon)  
 SOURCE Papio anubis  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Papio.  
 1 (bases 1 to 169904)  
 Akher, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,  
 Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,  
 Iaric, P., Lee-Lin, S.-O., Legaapi, R., Maduro, O.L., Maduro, V.B.,  
 Margulies, E.H., Masello, C., Maskeri, B., McQuinn, J.,  
 Pagittigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,  
 Reddy-Dugue, N., Schandier, K., Schneider, M.G., Sison, C.,  
 Stancirpop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,  
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 169904)  
 Green, E.D.  
 Direct Submission  
 Submitted (29-AUG-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 169904)  
 Green, E.D.  
 Direct Submission  
 Submitted (11-DEC-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 On Dec 11, 2002 this sequence version replaced gi:24308639.

----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: http://www.nisc.nih.gov  
 Contact: nisc\_zoo@hgri.nih.gov  
 ----- Project Information  
 Center project name: dli  
 Center clone name: 162P07  
 ----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 165859 bases at least Q40  
 Consensus quality: 16559 bases at least Q40  
 Consensus quality: 167094 bases at least Q20  
 Insert size: 14200; agarose-fp  
 Insert size: 169104; sum-of-coverage  
 Quality coverage: 14.6ix in Q20 bases; agarose-fp  
 Quality coverage: 12.27ix in Q20 bases; sum-of-coverage

----- NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 4138: contig of 4138 bp in length

4139 4238: gap of unknown length  
 \* 4239 10153: contig of 5915 bp in length  
 \* 10154 10253: gap of unknown length  
 \* 10254 18728: contig of 8475 bp in length  
 \* 18729 18828: gap of unknown length  
 \* 18829 32678: contig of 13850 bp in length  
 \* 32679 32778: gap of unknown length  
 \* 32779 54263: contig of 21485 bp in length  
 \* 54264 54363: gap of unknown length  
 \* 54364 76089: contig of 21726 bp in length  
 \* 76090 76189: gap of unknown length  
 \* 76190 101841: contig of 25652 bp in length  
 \* 101842 101941: gap of unknown length  
 \* 101942 123447: contig of 21506 bp in length  
 \* 123448 123547: gap of unknown length  
 \* 123548 169904: contig of 46357 bp in length.

FEATURES  
 source  
 1..169904  
 /organism="Papio anubis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9555"  
 /clone="RP41-162P7"  
 /clone\_11b="RP41"  
 1..4138  
 /note="assembly\_fragment  
 clone\_end:r7  
 vector\_side:left"  
 4239..10153  
 /note="assembly\_fragment"  
 10254..18728  
 /note="assembly\_fragment"  
 18829..32678  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:left"  
 32779..54263  
 /note="assembly\_fragment"  
 54364..76089  
 /note="assembly\_fragment"  
 76190..101841  
 /note="assembly\_fragment"  
 101942..123447  
 /note="assembly\_fragment"  
 123548..169904  
 /note="assembly\_fragment"

ORIGIN  
 Query Match 1.3%; Score 26; DB 2; Length 169904;  
 Best Local Similarity 100.0%; Pred. No. 0.04;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATAATTAATAATAGCTG 2036  
 |||  
 Db 102608 CCTACTAAATAATTAATAATAGCTG 102583

RESULT 92  
 AP002441 170404 bp DNA linear HTG 06-JUN-2000  
 LOCUS AP002441  
 DEFINITION Homo sapiens chromosome 18 clone RP11-702M18 map 18q22, WORKING  
 DRAFT SEQUENCE, 17 unordered pieces.

ACCESSION AP002441 GI:8307745  
 VERSION HTG: HTGS PHASE1: HTGS DRAFT.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 170404)  
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sekaki, Y.  
 Homo sapiens 170,404 genomic DNA of 18q22  
 JOURNAL Published Only in Database (2000)

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

2 (bases 1 to 170404)  
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,  
 Fujiyama, A., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.  
 Direct Substitution  
 Submitted (02-JUN-2000) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,  
 Japan (E-mail: hattori@gsc.riken.go.jp,  
 URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,  
 Fax: 81-42-778-9924)

## COMMENT

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

Project information

Center project name: RPI1-702M18

Center clone name: RPI1-702M18

Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 160385 bases at least Q40

Consensus quality: 165662 bases at least Q30

Insert size: 168804; sum-of-coverage

Quality coverage: 5.43x in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 39463 contig of 39463 bp in length
39564 69286 contig of 29723 bp in length
69387 89177 contig of 19791 bp in length
89278 107638 contig of 18616 bp in length
107739 117941 contig of 10703 bp in length
118042 125753 contig of 7712 bp in length
125854 131869 contig of 6016 bp in length
131970 137738 contig of 5769 bp in length
137839 143027 contig of 5189 bp in length
143128 147627 contig of 4399 bp in length
147627 151806 contig of 4180 bp in length
151907 155486 contig of 3580 bp in length
155486 159867 contig of 4281 bp in length
159867 163476 contig of 3509 bp in length
163476 165766 contig of 2190 bp in length
165766 168485 contig of 2619 bp in length
168485 170404 contig of 1819 bp in length.

NOTE: This is a 'working draft' sequence. It currently
consists of 17 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of 'N', but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 39463: contig of 39463 bp in length
39464 69286: gap of 100 bp
39564 69286: contig of 29723 bp in length
69287 89177: gap of 100 bp
69387 89177: contig of 19791 bp in length
89178 89278: gap of 100 bp
89279 107638: contig of 18616 bp in length
107639 107638: gap of 100 bp
107639 107638: gap of 100 bp
107739 117941: contig of 10703 bp in length
117942 118042: gap of 100 bp
118042 125753: contig of 7712 bp in length
125754 125854: gap of 100 bp
125854 131869: contig of 6016 bp in length

```

## FEATURES

## source

```

* 131870 131969: gap of 100 bp
* 131970 137738: contig of 5769 bp in length
* 137739 137838: gap of 100 bp
* 137839 143027: contig of 5189 bp in length
* 143028 143127: gap of 100 bp
* 143128 147526: contig of 4399 bp in length
* 147527 147627: gap of 100 bp
* 147627 151806: contig of 4180 bp in length
* 151807 151906: gap of 100 bp
* 151907 155486: contig of 3580 bp in length
* 155487 155586: gap of 100 bp
* 155587 159867: contig of 4281 bp in length
* 159868 159967: gap of 100 bp
* 159968 163476: contig of 3509 bp in length
* 163477 163576: gap of 100 bp
* 163577 165766: contig of 2190 bp in length
* 165767 168485: gap of 100 bp
* 168485 168585: contig of 2619 bp in length
* 168586 170404: gap of 100 bp
* 168586 170404: contig of 1819 bp in length.

Location/Qualifiers
1. 170404
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q22"
/clone="RPI1-702M18"
1. 39463
/note="assembly_fragment"
39564 69286
/note="assembly_fragment clone_end:SP6 vector_side:right"
69387 89177
/note="assembly_fragment"
89278 107638
/note="assembly_fragment"
107739 117941
/note="assembly_fragment"
118042 125753
/note="assembly_fragment"
125854 131869
/note="assembly_fragment"
131970 137738
/note="assembly_fragment"
137839 143027
/note="assembly_fragment"
143128 147526
/note="assembly_fragment"
147627 151806
/note="assembly_fragment"
151907 155486
/note="assembly_fragment"
155587 159867
/note="assembly_fragment"
159968 163476
/note="assembly_fragment"
163577 165766
/note="assembly_fragment clone_end:SP6 vector_side:left"
165867 168485
/note="assembly_fragment"
168586 170404
/note="assembly_fragment"

```

## ORIGIN

## Query Match

Best Local Similarity 1.3%; Score 26; DB 2; Length 170404;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

2011 CCTACTAAATATATTAATTCCTG 2036

## Db

66882 CCTACTAAATATATTAATTCCTG 66907

RESULT 93  
AC025132  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone RP11-201H6 map 11, WORKING DRAFT  
AC025132 170940 bp DNA linear HTG 24-AUG-2002  
SEQUENCE, 5 unordered pieces.  
AC025132  
AC025132.6 GI:13376953  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 11, clone RP11-201H6  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 170940)  
REFERENCE  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Boguslavsky, L., Meneses, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessilaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
REFERENCE  
AUTHORS 3 (bases 1 to 170940)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,  
Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,  
Collins, S., Collamore, A., Cooke, P., Dearellano, K., Dewar, K.,  
Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,  
Iliev, I., Johnson, R., Jones, C., Karatas, A., Lacroque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarthy, M.,  
McBwan, P., McKernan, K., McPheters, R., Meldrum, J., Meneses, L.,  
Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nobbu, C.,  
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,  
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,  
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,  
Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Straus, M., Subramanian, A., Talamas, J., Tessilaye, S.,  
Theodore, J., Travers, M., Trivis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
TITLE Direct Submission  
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Mar 18, 2001 this sequence version replaced g1:13357480.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L7837  
Center clone name: 201\_H6  
----- Summary Statistics  
Sequencing vector: M13; M77815; 3% of reads  
Sequencing vector: Plasmid; n/a; 97% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960721  
Consensus quality: 168745 bases at least Q40  
Consensus quality: 169718 bases at least Q30  
Consensus quality: 170224 bases at least Q20  
Insert size: 171000; agarose-fp  
Insert size: 170540; sum-of-ctnigs  
Quality coverage: 9.5 in Q20 bases; agarose-fp  
Quality coverage: 9.5 in Q20 bas.  
NOTE: This is a 'working draft' sequence. It currently  
consists of 5 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.  
1 12660: contig of 12660 bp in length  
12661 12760: gap of 100 bp  
12761 14981: contig of 2221 bp in length  
14982 15081: gap of 100 bp  
15082 101100: contig of 86019 bp in length  
101101 101200: gap of 100 bp  
101201 137473: contig of 36273 bp in length  
137474 137573: gap of 100 bp  
137574 170940: contig of 33367 bp in length.  
FEATURES  
source  
1..170940  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-201H6"  
/clone\_11b="RPCT-11 Human Male BAC"  
1..12660  
/note="assembly\_fragment  
clone end:SP6  
vector\_side:left"  
misc\_feature  
12761..14981  
/note="assembly\_fragment"  
misc\_feature  
15082..101100  
/note="assembly\_fragment"  
misc\_feature  
101201..137473  
/note="assembly\_fragment"  
misc\_feature  
137574..170940  
/note="assembly\_fragment  
clone end:T7  
vector\_side:right"  
ORIGIN  
Query Match 1.3%; Score 26; DB 2; Length 170940;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2011 CCTACTAAATATATAAATAGCTG 2036  
|||||  
Db 106255 CCTACTAAATATATAAATAGCTG 106280  
-----  
RESULT 94  
AC008937  
LOCUS AC008937 171480 bp DNA linear PRI 20-OCT-2001  
DEFINITION Homo sapiens chromosome 5 clone CTD-2310F14, complete sequence.  
AC008937

```

VERSION AC008937.7 GI:16303404
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 171480)
AUTHORS Doe Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 171480)
AUTHORS Doe Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 171480)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Oct 20, 2001 this sequence version replaced gi:13699460.
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.hgsc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
FEATURES
source
Location/Qualifiers
1. 171480
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2310F14"
ORIGIN
Query Match 1.3%; Score 26; DB 9; Length 171480;
Best Local Similarity 100.0%; Pred.No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2011 CCTACTAAATATATAAATTAGCTG 2036
Db 170925 CCTACTAAATATATAAATTAGCTG 170900
RESULT 95
AC026120 171998 bp DNA linear pri 22-OCT-2002
LOCUS Homo sapiens 12 BAC RP11-474N8 (Roswell Park Cancer Institute Human
AC026120 BAC library) complete sequence.
VERSION AC026120.33 GI:12642996
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 171998)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
1 (bases 1 to 171998)
Alshrock,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,B., Brown,M., Bryant,N.P., Buhay,C.,
Burich,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyne,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.U., Ding,Y., Dinn,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.H., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escocoto,M., Falls,T., Ferraguto,D.,
Flaegge,N., Ford,J., Foster,P., Franz,P., Gabriel,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guvvara,W.,
Gumarate,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homai,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
Karlsone,E., Kelly,S., Khan,Y., King,L., Kovach,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.,
Lewis,L., Li,J., Li,Z., Lichtharge,O., Lien,C., Liu,J., Liu,W.,
Louiseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,
Meador,M., Mei,G., Mercher,S., Metzger,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.J., Morgan,N.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newson,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenwo,S.,
Ogih,M., Okumura,G., Otagunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shooshart,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,P., Svatek,A., Taber,P., Tamezla,A.,
Tameria,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Wallington,S., Williams,G., Williamson,A.,
Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucheriapaci,R. and Gibbs,R.
Direct Submission
Unpublished
REFERENCE 2 (bases 1 to 171998)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 171998)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 171998)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE 6 (bases 1 to 171998)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 1, 2001 this sequence version replaced gi:12621199.
COMMENT INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.
ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

```





Yiel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (25-MAR-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 172192)

Anderson, S., Barna, N., Bastien, C., Lander, E., Ali, A., Allen, N.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Cooke, P., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacombe, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., MacDonald, P., Major, J., Marguis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trifillo, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission

Submitted (06-SEP-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 172192)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhassir, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Cooke, P., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacombe, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., MacDonald, P., Major, J., Marguis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trifillo, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission

Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Dec 11, 2001 this sequence version replaced gi:15451709.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/MN/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: l13190

Center clone name: 702\_M\_18

Location/Qualifiers

1. 172192

/organism="Homo sapiens"

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-702M18"
/clone_1ib="RP11-702M18"
/complement(2..860)
/rpc_family="L1M4"
764
/note="Is definitely AGA, not AAA."
complement(1019..1247)
/rpc_family="L1M4"
1570..1574
/note="30 qual SGL region"
2118..2189
/rpc_family="L1M4"
complement(1207..2519)
/rpc_family="L1M4"
2888..2937
/rpc_family="L1M4"
complement(3611..3900)
/rpc_family="L1M4"
complement(4278..6251)
/rpc_family="L1M4"
6252..6959
/rpc_family="L1M4"
6957..7161
/rpc_family="L1M4"
7180..7218
/rpc_family="L1M4"
7219..7845
/rpc_family="L1M4"
8709..8743
/rpc_family="L1M4"
9473..9785
/rpc_family="L1M4"
10407..10984
/rpc_family="L1M4"
11025..11183
/rpc_family="L1M4"
11231..11331
/rpc_family="L1M4"
12274..12543
/rpc_family="L1M4"
13718..13980
/rpc_family="L1M4"
14899..14929
/rpc_family="L1M4"
complement(15279..15486)
/rpc_family="L1M4"
complement(15587..15732)
/rpc_family="L1M4"
complement(15922..16279)
/rpc_family="L1M4"
complement(16313..16391)
/rpc_family="L1M4"
16402..17369
/rpc_family="L1M4"
17429..17451
/rpc_family="L1M4"
17455..17490
/rpc_family="L1M4"
17490..17534
/rpc_family="L1M4"
complement(17724..18007)
/rpc_family="L1M4"
complement(19261..19515)
/rpc_family="L1M4"
complement(19610..19754)
/rpc_family="L1M4"
19769..20120
/rpc_family="L1M4"
20121..20424

```

```
repeat_region      /rpt_family="ALUy"
                    complement(20443, .20813)
repeat_region      /rpt_family="L1ME1"
                    20814..20858
repeat_region      /rpt_family="(CA)n"
                    complement(20859, .21182)
repeat_region      /rpt_family="L1ME1"
                    complement(21242, .21401)
repeat_region      /rpt_family="L1M4"
                    21402..21501
repeat_region      /rpt_family="L1MB8"
                    complement(21502, .21710)
repeat_region      /rpt_family="L1M4"
                    complement(21709, .21938)
repeat_region      /rpt_family="L1MA8"
                    complement(21946, .22162)
repeat_region      /rpt_family="L1M4"
                    complement(22168, .22238)
repeat_region      /rpt_family="LTR49-int"
                    complement(22260, .22756)
repeat_region      /rpt_family="L1M4"
```

```
Query Match      1.3%; Score 26; DB 9; Length 172192;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2011 CCTACTAAATAATTAATAATGCTG 2036
Db      2406 CCTACTAAATAATTAATAATGCTG 2381
```

```
RESULT 97
AP001808      172357 bp      DNA      linear      HTG 30-MAY-2000
LOCUS      Homo sapiens chromosome 18 clone RP11-699G7 map 18q12, WORKING
DEFINITION      DRAFT SEQUENCE. 24 unordered pieces.
ACCESSION      AP001808
VERSION      AP001808.2 GI:8117479
KEYWORDS      HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 172357)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 172,357 genomic DNA of 18q12
2 (bases 1 to 172357)
Published Only in DataBase (2000)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submision
Submitted (14-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Katsushika Univ., 1-15-1 Katsushika, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 30, 2000 this sequence version replaced gi:7594895.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraFt18
Center clone name: RP11-699G7
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amcramham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 157675 bases at least Q40
Consensus quality: 165001 bases at least Q30
```

Consensus quality: 168122 bases at least Q20  
Insert size: 170057; sum-of-contigs  
Quality coverage: 4.83x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1      20111 contig of 20111 bp in length
1      20212 35529 contig of 15318 bp in length
35630 47492 contig of 11863 bp in length
47593 59240 contig of 11648 bp in length
59341 74437 contig of 15097 bp in length
74538 84647 contig of 10110 bp in length
84748 93348 contig of 8601 bp in length
93449 103418 contig of 9970 bp in length
103519 110404 contig of 6785 bp in length
110404 116483 contig of 6080 bp in length
116483 123488 contig of 6905 bp in length
123488 123588 gap of 100 bp
123588 129188 contig of 5600 bp in length
129188 129288 gap of 100 bp
129288 133859 contig of 4571 bp in length
133859 133960 gap of 100 bp
133960 139541 contig of 5582 bp in length
139541 139641 gap of 100 bp
139641 145042 contig of 5401 bp in length
```

```

* 145043 145142: gap of 100 bp
* 145143 150084: contig of 4942 bp in length
* 150085 150184: gap of 100 bp
* 150185 154055: contig of 3871 bp in length
* 154056 154155: gap of 100 bp
* 154156 158143: contig of 3988 bp in length
* 158144 158243: gap of 100 bp
* 158244 161205: contig of 2962 bp in length
* 161206 161305: gap of 100 bp
* 161306 164028: contig of 2723 bp in length
* 164029 164128: gap of 100 bp
* 164129 165241: contig of 1113 bp in length
* 165242 165341: gap of 100 bp
* 165342 168406: contig of 3065 bp in length
* 168407 168506: gap of 100 bp
* 168507 170188: contig of 1682 bp in length
* 170189 170289: gap of 100 bp
* 170289 172357: contig of 2069 bp in length.

FEATURES
    source
        1..172357
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="18"
            /map="18q12"
            /clone="RP11-699G7"
            1..20111
                /note="assembly_fragment"
            20112..35529
                /note="assembly_fragment"
            35530..47492
                /note="assembly_fragment"
            47593..59240
                /note="assembly_fragment"
            59341..74437
                /note="assembly_fragment"
            74538..84647
                /note="assembly_fragment"
            84748..93348
                /note="assembly_fragment"
            93449..103418
                /note="assembly_fragment"
            103519..110303
                /note="assembly_fragment"
            110404..116483
                /note="assembly_fragment"
            116584..123488
                /note="assembly_fragment"
            123589..129188
                /note="assembly_fragment"
            129289..133859
                /note="assembly_fragment"
            133960..139541
                /note="assembly_fragment"
            139642..145042
                /note="assembly_fragment"
            145143..150084
                /note="assembly_fragment"
            150185..154055
                /note="assembly_fragment"
            154156..158143
                /note="assembly_fragment"
            158244..161205
                /note="assembly_fragment"
            161306..164028
                /note="assembly_fragment"
            164129..165241
                /note="assembly_fragment"
            165342..168406
                /note="assembly_fragment"
            168507..170188
                /note="assembly_fragment"
            170289..172357
                /note="assembly_fragment"

```

```

ORIGIN
    /note="assembly_fragment"

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 172357;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAATATTAATTTGCTG 2036
DB 88776 CCTACTAAATATTAATTTAGCTG 88801

RESULT 98
AC027063
LOCUS
DEFINITION
AC027063 Homo sapiens clone RP11-605A10, WORKING DRAFT SEQUENCE, 21
uncloned pieces.
ACCESSION
AC027063
VERSION
AC027063.3 GI:8077002
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 173735)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-605A10
Unpublished
2 (bases 1 to 173735)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collins, R., Cooke, P., Dearrell, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A.,
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McSheeters, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (26-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7637290.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 605_A_10
Center clone name: L7466
----- Summary Statistics
Sequencing vector: M13; 877815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 163013 bases at least Q40
Consensus quality: 168440 bases at least Q30
Consensus quality: 170590 bases at least Q20
Insert size: 170000; agarose-fp

```

Insert size: 171735; sum-of-contigs  
 Quality coverage: 4.6 in Q20 bases; agarose-fp  
 Quality coverage: 4.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1198: contig of 1198 bp in length
* 1199: gap of 100 bp
* 1298: gap of 100 bp
* 1299: gap of 1506 bp in length
* 1299: gap of 100 bp
* 2805: gap of 100 bp
* 2904: gap of 100 bp
* 2905: gap of 1997 bp in length
* 4901: contig of 1997 bp in length
* 4902: gap of 100 bp
* 5001: gap of 100 bp
* 5002: contig of 3341 bp in length
* 8342: gap of 100 bp
* 8343: gap of 100 bp
* 8442: gap of 100 bp
* 8443: contig of 2848 bp in length
* 11290: contig of 2848 bp in length
* 11291: gap of 100 bp
* 11298: gap of 3896 bp in length
* 11391: gap of 100 bp
* 15286: gap of 100 bp
* 15287: gap of 100 bp
* 15387: contig of 3967 bp in length
* 19353: contig of 3967 bp in length
* 19354: gap of 100 bp
* 19454: gap of 100 bp
* 24735: contig of 5282 bp in length
* 24736: gap of 100 bp
* 24835: gap of 100 bp
* 30442: contig of 5607 bp in length
* 30443: gap of 100 bp
* 30542: gap of 100 bp
* 38141: contig of 7599 bp in length
* 38142: gap of 100 bp
* 38241: gap of 100 bp
* 38242: contig of 7015 bp in length
* 45256: contig of 7015 bp in length
* 45257: gap of 100 bp
* 45357: contig of 7031 bp in length
* 52387: contig of 7031 bp in length
* 52388: gap of 100 bp
* 52487: gap of 100 bp
* 60034: contig of 7547 bp in length
* 60134: gap of 100 bp
* 60135: gap of 100 bp
* 68588: contig of 8454 bp in length
* 68689: gap of 100 bp
* 68689: gap of 100 bp
* 76018: contig of 7330 bp in length
* 76019: gap of 100 bp
* 76118: gap of 100 bp
* 86511: contig of 10393 bp in length
* 86512: gap of 100 bp
* 86512: gap of 100 bp
* 95770: contig of 9159 bp in length
* 95771: gap of 100 bp
* 95870: gap of 100 bp
* 95871: contig of 11283 bp in length
* 107153: contig of 11283 bp in length
* 107154: gap of 100 bp
* 107253: gap of 100 bp
* 107254: contig of 18242 bp in length
* 125495: contig of 18242 bp in length
* 125496: gap of 100 bp
* 125595: gap of 100 bp
* 144262: contig of 18667 bp in length
* 144263: gap of 100 bp
* 144363: contig of 29373 bp in length.

```

## FEATURES

## source

```

1. .173735
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-605A10"
/clone_11b="RP11-605A10"
/clone_11b="RP11-605A10"
1. .1198
/misc_feature
/note="assembly_fragment"
1299. .2804
/misc_feature
/note="assembly_fragment"
2905. .4901
/misc_feature
/note="assembly_fragment"
5002. .8342
/misc_feature
/note="assembly_fragment"
8443. .11290
/misc_feature
/note="assembly_fragment"
11391. .15286
/misc_feature
/note="assembly_fragment"
15387. .19353
/misc_feature

```

## ORIGIN

```

misc_feature
/note="assembly_fragment"
19454. .24735
/note="assembly_fragment"
24836. .30442
/note="assembly_fragment"
30543. .38141
/note="assembly_fragment"
38242. .45256
/note="assembly_fragment"
45357. .52387
/note="assembly_fragment"
52487. .60034
/note="assembly_fragment"
60135. .68588
/note="assembly_fragment"
68689. .76018
/note="assembly_fragment"
76119. .86511
/note="assembly_fragment"
86612. .95770
/note="assembly_fragment"
95871. .107153
/note="assembly_fragment"
107254. .125495
/note="assembly_fragment"
125596. .144262
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
144363. .173735
/note="assembly_fragment"

```

```

Query Match 1.3%; Score 26; DB 2; Length 173735;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2011 CCTACTAAATATATAAATTAGCTG 2036
Db 116074 CCTACTAAATATATAAATTAGCTG 116099

```

```

RESULT 99
AL392084 173942 bp DNA linear PRI 04-APR-2001
LOCUS AL392084
DEFINITION Human DNA sequence from clone RP11-576E23 on chromosome 9 Contains
ACCESSION AL392084
VERSION AL392084.6 GI:11611388
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 173942)
REFERENCE
1 Williams, S.
Direct Submision
Submitted (06-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
UK
Requesters: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11557944.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMBEP; Information
on the WORMBEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormep This sequence
was generated from part of bacterial clone contigs of human

```

chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>  
RP11-576E23 is from the library RP11-1.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/dacpac/home.htm>  
VECTOR: pBACE3.6

This sequence is the entire insert of clone RP11-576E23. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## FEATURES

```
source
    location/Qualifiers
    1..173942
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="9"
    /clone="RP11-576E23"
    /clone_lib="RP11-1.2"
    /complement(149..1020)
    /note="match: GSS: Em:AQ745338"
    /complement(532..1038)
    /note="match: GSS: Em:AQ668228"
    867..923
    /note="L1PBA repeat: matches -1217..-1164 of consensus"
    1587..1950
    /note="MSTA repeat: matches 1..409 of consensus"
    2300..3217
    /note="L1MB8 repeat: matches 4904..5800 of consensus"
    3218..3577
    /note="TRHIC repeat: matches 1..371 of consensus"
    3578..3760
    /note="L1MB8 repeat: matches 5800..5977 of consensus"
    3761..4015
    /note="L1PAL3 repeat: matches 5900..6155 of consensus"
    4016..4160
    /note="L1MB8 repeat: matches 5977..6173 of consensus"
    5458..5696
    /note="MIR repeat: matches 8..250 of consensus"
    /complement(5527..5839)
    /note="match: GSS: Em:B17381"
    5740..5790
    /note="17 copies 3 mer atc 94% conserved"
    6035..6193
    /note="MIR repeat: matches 81..242 of consensus"
    7718..7785
    /note="L2 repeat: matches 2643..2710 of consensus"
    10106..10781
    /note="338 copies 2 mer aa 60% conserved"
    10802..11662
    /note="L1PBB repeat: matches 937..1770 of consensus"
    11734..14963
    /note="L1PBB repeat: matches 2919..6150 of consensus"
    15390..15484
    /note="match: GSS: Em:AQ260394"
    15390..15484
    /note="MIR repeat: matches 31..128 of consensus"
    15404..15823
    /note="match: GSS: Em:AQ275766"
    15647..15768
    /note="MERSA repeat: matches 55..189 of consensus"
    /complement(17237..17864)
    /note="match: GSS: Em:AQ260131"
    /complement(17395..17859)
    /note="match: GSS: Em:AQ260391"
    17431..17499
    /note="MERSB repeat: matches 1..69 of consensus"
    17492..17546
    /note="L2 repeat: matches 2683..2745 of consensus"

repeat_region
    17669..17919
    /note="MIR repeat: matches 33..262 of consensus"
    18447..19191
    /note="L1MB8 repeat: matches 5558..6307 of consensus"
    19244..19545
    /note="AluX repeat: matches 1..303 of consensus"
    19763..19844
    /note="MIR repeat: matches 117..206 of consensus"
    19885..20040
    /note="FRAM repeat: matches 1..157 of consensus"
    20358..20909
    /note="MERSA repeat: matches 1..605 of consensus"
    21058..21129
    /note="MERSA repeat: matches 117..188 of consensus"
    21631..21685
    /note="MIR repeat: matches 172..225 of consensus"
    21749..21829
    /note="MIR repeat: matches 71..153 of consensus"
    22038..22209
    /note="MIR repeat: matches 70..248 of consensus"
    22361..22447
    /note="MIR repeat: matches 127..213 of consensus"
    22738..22956
    /note="L2 repeat: matches 2537..2749 of consensus"
    23078..23193
    /note="L2 repeat: matches 2613..2750 of consensus"
    23395..23518
    /note="L2 repeat: matches 2625..2749 of consensus"
    23935..24138
    /note="MIR repeat: matches 27..258 of consensus"
    24156..24576
    /note="L2 repeat: matches 2310..2708 of consensus"
    /complement(24226..24702)
    /note="match: GSS: Em:B57730"
    /complement(24294..24753)
    /note="match: GSS: Em:AQ220752"
    /complement(24371..24753)
    /note="match: GSS: Em:AQ077990"
    24699..24772
    /note="MERSA repeat: matches 33..112 of consensus"
    25271..25578
    /note="AluSc repeat: matches 1..309 of consensus"
    25840..26256
    /note="MERS7 repeat: matches 18..509 of consensus"
    26400..26712
    /note="AluX repeat: matches 2..312 of consensus"
    26717..27154
    /note="L1MCA repeat: matches 7488..7977 of consensus"
    27177..28029
    /note="L1MCA repeat: matches 6684..7737 of consensus"
    28030..28180
    /note="L1MCA repeat: matches 7585..7735 of consensus"
    29315..29405
    /note="MIR repeat: matches 41..144 of consensus"
    29469..29741
    /note="Alu repeat: matches 38..310 of consensus"
    29891..30238
    /note="L1R16C repeat: matches 1..353 of consensus"
    30432..30511
    /note="MLT17 repeat: matches 120..200 of consensus"
    31383..31612
    /note="MIR repeat: matches 3..251 of consensus"
    31626..32191
    /note="match: GSS: Em:AQ491560"
    32271..32493
    /note="MIR repeat: matches 21..254 of consensus"
    32529..32572
    /note="11 copies 4 mer caca 77% conserved"
    32916..33070
    /note="MIR repeat: matches 20..183 of consensus"
    33086..33173
    /note="L1PA6 repeat: matches 6055..6141 of consensus"
    33279..33784
```

```

repeat_region /note="L1P6 repeat: matches 5572. .6075 of consensus"
33933. .34131
repeat_region /note="L2 repeat: matches 2080. .2284 of consensus"
34389. .34430
repeat_region /note="21 copies 2 mer tc 92% conserved"
34391. .34430
repeat_region /note="10 copies 4 mer tctc 95% conserved"
34457. .34541
repeat_region /note="L2 repeat: matches 2618. .2703 of consensus"
35679. .35817
repeat_region /note="L2 repeat: matches 5708. .5851 of consensus"
36131. .36531
repeat_region /note="L1P6 repeat: matches 21. .449 of consensus"
36131. .36531
misc_feature /note="match: GSS: Em:AQ137035"
36890. .37438
repeat_region /note="L2 repeat: matches 1991. .2564 of consensus"
37439. .37744
repeat_region /note="L1P6 repeat: matches 1. .299 of consensus"
37745. .37879
repeat_region /note="L2 repeat: matches 2564. .2684 of consensus"
38245. .38306
repeat_region /note="L2 repeat: matches 78. .141 of consensus"
38321. .38441
repeat_region /note="L2 repeat: matches 17. .164 of consensus"
38321. .38441

```

```

Query Match 1.3%; Score 26; DB 9; Length 173942;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2011 CCTACTAAATATATAAATTAGCTG 2036
DB 26513 CCTACTAAATATATAAATTAGCTG 26538

```

```

RESULT 100
AL450337 173987 bp DNA linear PRI 15-NOV-2001
LOCUS Human DNA sequence from clone RP11-51B10 on chromosome 10, complete
DEFINITION sequence.
ACCESSION AL450337
VERSION AL450337.24 GI:16972991
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 173987)

```

```

REFERENCE Direct Submission
AUTHORS Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
TITLE Cambridge, CB10 1SA, UK. E-mail enquiries:
JOURNAL humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Nov 16, 2001 this sequence version replaced GI:16944060.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em; EMBL; SW;
SWISSPROT; Tr; TrEMBL; Wp; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human

```

chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10>

RP11-51B10 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-51B10. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true left end of clone RP11-51B10 is at 1 in this sequence. The true left end of clone RP11-297A16 is at 171988 in this sequence. The true right end of clone RP11-363N3 is at 123830 in this sequence.

```

FEATURES
source 1. .173987
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-51B10"
/clone_1fb="RP11-11.1"
13053. .13117
/misc_feature /note="sequence from uni-directional dGTP big dye
terminator reads only"
118362. .118392
/misc_feature /note="Sequence from uni-directional dGTP big dye
terminator reads only."

```

```

Query Match 1.3%; Score 26; DB 9; Length 173987;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2011 CCTACTAAATATATAAATTAGCTG 2036
DB 106926 CCTACTAAATATATAAATTAGCTG 106951

```

Search completed: January 4, 2005, 20:59:13  
Job time : 8405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using bw model

Run on: January 4, 2005, 17:57:14 ; Search time 924 Seconds  
(without alignments)  
11566.906 Million cell updates/sec

Title: US-09-508-658B-1

Perfect score: 2036

Sequence: 1 agaccggggagagacggcg99.....aaaatacaaaatagctg 2036

Scoring table: ~~OLIGOJMC~~  
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database :

1: N\_Geneseq\_23sep04:\*  
2: geneeqn1980s:\*  
3: geneeqn2000s:\*  
4: geneeqn2001as:\*  
5: geneeqn2001bs:\*  
6: geneeqn2002as:\*  
7: geneeqn2002bs:\*  
8: geneeqn2003as:\*  
9: geneeqn2003bs:\*  
10: geneeqn2003cs:\*  
11: geneeqn2003ds:\*  
12: geneeqn2004s:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2036	100.0	2036	2	AAx26936 cDNA enco
2	1969	96.7	2245	2	AAx58605 Human aut
3	1025	50.3	1545	2	AAx26937 cDNA enco
4	692	34.0	1463	2	AAx26938 cDNA enco
5	178	8.7	556	12	ACH67354 Human gen
6	175	8.6	175	12	ACH81059 Human gen
7	150	7.4	530	12	ACH8194 Human gen
8	146	7.2	146	12	ACH81896 Human gen
9	60	2.9	60	6	ABN38615 Human sp1
10	32	1.6	1656	2	AAx58606 Mouse aut
11	26	1.3	9742	4	AAx42015 Genomic B
12	25	1.2	95	4	AAx71122 Human imm
13	25	1.2	113	5	ABx16553 Human ner
14	25	1.2	147	4	AAx74219 Human imm
15	25	1.2	147	4	AAx74218 Human imm
16	25	1.2	172	3	AAx12190 Human sec
17	25	1.2	213	4	AAx104285 Human rep
18	25	1.2	239	3	AAx20108 Human sec
19	25	1.2	268	3	AAx57008 Human col
20	25	1.2	268	6	ABT12430 Orestes B
21	25	1.2	268	10	ACD91724 Human col

C 22	25	1.2	274	4	AAK83336 Human imm
C 23	25	1.2	282	4	AAK75842 Human imm
C 24	25	1.2	282	6	AB177830 Human ova
C 25	25	1.2	292	2	AAV88195 EST clone
C 26	25	1.2	302	4	AAK76822 Human imm
C 27	25	1.2	357	5	AAx65644 Human imm
C 28	25	1.2	375	4	AAK78393 Human imm
C 29	25	1.2	375	4	AAK78396 Human imm
C 30	25	1.2	389	4	AAK81129 Human imm
C 31	25	1.2	389	4	AAK81130 Human imm
C 32	25	1.2	401	4	AAK96262 Human neu
C 33	25	1.2	401	4	AAK97755 Human neu
C 34	25	1.2	401	6	ABT01032 Human neu
C 35	25	1.2	403	5	ABT02525 Human neu
C 36	25	1.2	403	5	ABV57504 Human pro
C 37	25	1.2	411	2	AAx05173 Human MSH
C 38	25	1.2	427	5	ABV17272 Human pro
C 39	25	1.2	434	9	ACH47498 Human inf
C 40	25	1.2	447	4	AAH13419 Human can
C 41	25	1.2	447	6	ABN61847 Human can
C 42	25	1.2	453	4	AAK62554 Human imm
C 43	25	1.2	467	5	ABV47067 Human pro
C 44	25	1.2	472	6	AB183732 Human ova
C 45	25	1.2	474	9	ACH15457 Human adu
C 46	25	1.2	486	9	ACH28081 Human adu
C 47	25	1.2	488	9	ACH47229 Human inf
C 48	25	1.2	496	4	AAH13219 Human CDN
C 49	25	1.2	503	10	ADF90604 Human hep
C 50	25	1.2	520	4	AAH15788 Human bre
C 51	25	1.2	541	5	AAx94242 Primer sp
C 52	25	1.2	550	6	ABN60792 Human can
C 53	25	1.2	553	12	ADN12564 Human pro
C 54	25	1.2	555	4	AAH10441 Human CDN
C 55	25	1.2	582	4	AAH10596 Human CDN
C 56	25	1.2	585	5	ABV51377 Human pro
C 57	25	1.2	590	6	ABN60118 Human can
C 58	25	1.2	605	5	ABV58013 Human pro
C 59	25	1.2	613	8	AAH36536 Human mus
C 60	25	1.2	613	8	ABX59524 cDNA enco
C 61	25	1.2	613	12	ADJ30274 Human mus
C 62	25	1.2	684	5	ADU42725 Human ova
C 63	25	1.2	685	3	ADN00585 Alternati
C 64	25	1.2	692	4	AAH04652 Human rep
C 65	25	1.2	692	4	AAH04651 Human rep
C 66	25	1.2	700	4	AAH92601 Human inf
C 67	25	1.2	700	4	AAH92602 Human inf
C 68	25	1.2	703	5	ABV42610 Human pro
C 69	25	1.2	742	5	ADL62169 Human ova
C 70	25	1.2	743	9	ADB82582 Human CDN
C 71	25	1.2	749	4	AAH24632 Human bre
C 72	25	1.2	781	4	AAH195858 Human neu
C 73	25	1.2	816	6	ABQ89641 Human pro
C 74	25	1.2	852	6	ABX92120 Lung spec
C 75	25	1.2	1001	3	AAx57995 Atrichidon
C 76	25	1.2	1001	3	AAx57996 Atrichidon
C 77	25	1.2	1055	5	AAH94439 Human foe
C 78	25	1.2	1055	5	AAH94410 Human foe
C 79	25	1.2	1107	5	ABV28056 Human rep
C 80	25	1.2	1229	4	AAH03604 Human ova
C 81	25	1.2	1229	4	ABAO7805 Human imm
C 82	25	1.2	1255	4	AAK81121 Human imm
C 83	25	1.2	1265	5	ABV25475 Human pro
C 84	25	1.2	1289	4	AAH18060 Human CDN
C 85	25	1.2	1382	8	ABZ71641 Breast sp
C 86	25	1.2	1382	12	ADF85883 Human bre
C 87	25	1.2	1387	4	AAH12601 Human pro
C 88	25	1.2	1801	4	AAH17545 Human CDN
C 89	25	1.2	1822	2	AAV34248 Human sec
C 90	25	1.2	1822	8	ACD08119 cDNA enco
C 91	25	1.2	1859	4	AAx26824 Human CDN
C 92	25	1.2	1861	4	AAx28995 Genomic B
C 93	25	1.2	1861	4	AAx31660 Genomic B
C 94	25	1.2	1861	6	ABT07830 Novel hum

95	1.2	1869	4	ABA06374	AbA06374 Human cDN	168	25	1.2	2213	10	ACD68677	AcD68677 Novel hum
96	1.2	1869	6	ABV83719	AbV83719 Human pol	169	25	1.2	2213	12	ADC48872	AdC48872 Novel hum
97	1.2	1880	4	AAH15709	AaH15709 Human cDN	170	25	1.2	2213	12	ADE21043	AdE21043 Novel hum
98	1.2	1947	4	AAH16703	AaH16703 Human cDN	171	25	1.2	2213	12	ADE05887	AdE05887 Novel hum
99	1.2	1993	10	ACA64683	AcA64683 DNA encod	172	25	1.2	2213	12	ADD75116	AdD75116 Novel hum
C 100	1.2	1993	10	ACA64684	AcA64684 DNA encod	173	25	1.2	2213	12	ADD75862	AdD75862 Novel hum
C 101	1.2	1994	10	ACA64687	AcA64687 DNA encod	174	25	1.2	2213	12	ADD85094	AdD85094 Novel hum
C 102	1.2	1994	10	ACA64678	AcA64678 DNA encod	175	25	1.2	2213	12	ADD86920	AdD86920 Novel hum
C 103	1.2	1994	10	ACA64681	AcA64681 DNA encod	176	25	1.2	2213	12	ADD82097	AdD82097 Novel hum
C 104	1.2	1994	10	ACA64685	AcA64685 DNA encod	177	25	1.2	2213	12	ADE33094	AdE33094 Novel hum
C 105	1.2	1994	10	ACA64686	AcA64686 DNA encod	178	25	1.2	2213	12	ADE05641	AdE05641 Novel hum
C 106	1.2	1994	10	ACA64682	AcA64682 DNA encod	179	25	1.2	2213	12	ADD73626	AdD73626 Novel hum
C 107	1.2	2081	10	ADE08852	AdE08852 Novel DNA	180	25	1.2	2213	12	ADD78466	AdD78466 Novel hum
C 108	1.2	2080	5	ABV22219	AbV22219 Human pro	181	25	1.2	2213	12	ADE41439	AdE41439 Novel hum
C 109	1.2	2180	3	AAV22219	AaV22219 Human sec	182	25	1.2	2213	12	ADE21289	AdE21289 Novel hum
C 110	1.2	2213	3	AAV77666	AaV77666 Human PRO	183	25	1.2	2213	12	ADD77404	AdD77404 Novel hum
C 111	1.2	2213	3	AAV37058	AaV37058 Human PRO	184	25	1.2	2213	12	ADE20551	AdE20551 Novel hum
C 112	1.2	2213	4	AAV54294	AaV54294 DNA encod	185	25	1.2	2213	12	ADD75616	AdD75616 Novel hum
C 113	1.2	2213	6	ABK33620	AbK33620 cDNA enco	186	25	1.2	2213	12	ADD74132	AdD74132 Novel hum
C 114	1.2	2213	6	ABL88166	AbL88166 Human PRO	187	25	1.2	2213	12	ADD74378	AdD74378 Novel hum
C 115	1.2	2213	8	ABL95655	AbL95655 Human ang	188	25	1.2	2213	12	ADD76108	AdD76108 Novel hum
C 116	1.2	2213	8	ACA68581	AcA68581 Novel hum	189	25	1.2	2213	12	ADD85600	AdD85600 Novel hum
C 117	1.2	2213	9	ABT44310	AbT44310 Human PRO	190	25	1.2	2213	12	ADE05149	AdE05149 Novel hum
C 118	1.2	2213	9	ACD68331	AcD68331 Novel hum	191	25	1.2	2213	12	ADD75362	AdD75362 Novel hum
C 119	1.2	2213	9	ABT44593	AbT44593 Human PRO	192	25	1.2	2213	12	ADD75906	AdD75906 Novel hum
C 120	1.2	2213	9	ACD82260	AcD82260 Human sec	193	25	1.2	2213	12	ADD86674	AdD86674 Novel hum
C 121	1.2	2213	9	ACH04433	AcH04433 Human cDN	194	25	1.2	2213	12	ADD78142	AdD78142 Novel hum
C 122	1.2	2213	9	ACD67977	AcD67977 Novel hum	195	25	1.2	2213	12	ADD77650	AdD77650 Novel hum
C 123	1.2	2213	9	ABT43966	AbT43966 Human mem	196	25	1.2	2213	12	ADD77896	AdD77896 Novel hum
C 124	1.2	2213	9	ADB883659	AdB883659 Novel hum	197	25	1.2	2213	12	ADD85354	AdD85354 Novel hum
C 125	1.2	2213	9	ADB880765	AdB880765 Novel hum	198	25	1.2	2213	12	ADD73886	AdD73886 Novel hum
C 126	1.2	2213	9	ADB73306	AdB73306 Novel hum	199	25	1.2	2213	12	ADD74624	AdD74624 Novel hum
C 127	1.2	2213	9	ADB78388	AdB78388 Novel hum	200	25	1.2	2213	12	ADD77152	AdD77152 Novel hum
C 128	1.2	2213	10	ADB85036	AdB85036 Novel hum							
C 129	1.2	2213	10	ADB78142	AdB78142 Novel hum							
C 130	1.2	2213	10	ADB87208	AdB87208 Human PRO							
C 131	1.2	2213	10	ADB84790	AdB84790 Human PRO							
C 132	1.2	2213	10	ADB883905	AdB883905 Novel hum							
C 133	1.2	2213	10	ADB73060	AdB73060 Novel hum							
C 134	1.2	2213	10	ADC17998	AdC17998 Human PRO							
C 135	1.2	2213	10	ADC36898	AdC36898 Human PRO							
C 136	1.2	2213	10	ADC21888	AdC21888 Human PRO							
C 137	1.2	2213	10	ADC49919	AdC49919 Novel hum							
C 138	1.2	2213	10	ADC499118	AdC499118 Novel hum							
C 139	1.2	2213	10	ADC49635	AdC49635 Novel hum							
C 140	1.2	2213	10	ADC47496	AdC47496 Novel hum							
C 141	1.2	2213	10	ADC47241	AdC47241 Novel hum							
C 142	1.2	2213	10	ADC78116	AdC78116 Novel hum							
C 143	1.2	2213	10	ADD06351	AdD06351 Novel hum							
C 144	1.2	2213	10	ADD10478	AdD10478 Human sec							
C 145	1.2	2213	10	ADC7870	AdC7870 Novel hum							
C 146	1.2	2213	10	ADD11438	AdD11438 Human sec							
C 147	1.2	2213	10	ADD50833	AdD50833 Novel hum							
C 148	1.2	2213	10	ADD51079	AdD51079 Novel hum							
C 149	1.2	2213	10	ADD70644	AdD70644 Human cDN							
C 150	1.2	2213	10	ADD39721	AdD39721 Human cDN							
C 151	1.2	2213	10	ADD70167	AdD70167 Human cDN							
C 152	1.2	2213	10	ADD37231	AdD37231 Human sec							
C 153	1.2	2213	10	ADD38288	AdD38288 Human cDN							
C 154	1.2	2213	10	ADD39244	AdD39244 Human cDN							
C 155	1.2	2213	10	ADD50560	AdD50560 Human PRO							
C 156	1.2	2213	10	ADD50314	AdD50314 Human PRO							
C 157	1.2	2213	10	ADD38767	AdD38767 Human cDN							
C 158	1.2	2213	10	ADD40198	AdD40198 Human cDN							
C 159	1.2	2213	10	ADD51325	AdD51325 Novel hum							
C 160	1.2	2213	10	AD550419	AdE550419 Human cDN							
C 161	1.2	2213	10	AD520031	AdE20031 Human cDN							
C 162	1.2	2213	10	AD549942	AdE49942 Human cDN							
C 163	1.2	2213	10	AD521500	AdE21500 Human cDN							
C 164	1.2	2213	10	AD529925	AdE29925 Human cDN							
C 165	1.2	2213	10	AD555818	AdE55818 Human cDN							
C 166	1.2	2213	10	ADH99322	AdH99322 Human cDN							
C 167	1.2	2213	10	ACA66925	AcA66925 cDNA enco							

## ALIGNMENTS

RESULT 1	
ID AAX26936	strand: DNA, 2036 BP.
AAAX26936;	
24-JUN-1999	(first entry)
CDNA encoding a human autoimmune regulator-1 (AIR-1) protein.	
Autoimmune regulator-1; AIR-1; immune maturation; immune response;	
disease; autoimmune polyendocrinopathy candidiasis ectodermal dystrophy;	
APECED; autoimmune polyglandular syndrome type I; APS I; ss.	
Homo sapiens.	
Key	Location/Qualifiers
FT CDS	137..1774
FT	/*tag= a
FT	/product= "AIR-1"
OS	WO9915559-A1.
PD	01-APR-1999.
PR	23-SEP-1998; 98WO-FI000749.
XX	23-SEP-1997; 97FI-00003762.
XX	(FIM-) FINNISH IMMUNOTECHNOLOGY LTD.
XX	Krohn K, Helio M, Peterson P, Scott H, Antonarakis S, Lalioti M;
PI	Shimizu N, Kudoh J;
XX	WPI; 1999-244390/20.



DR P-PSDB; AAY01712.  
XX Autoimmune regulator 1 (AIR1) DNA sequence.  
PT Claim 4; Page 21-24; 59pp; English.  
XX  
XX  
XX The present sequence encodes an autoimmune regulator-1 (AIR-1) protein.  
CC The AIR polypeptides and polynucleotides can be used in methods for the  
CC diagnosis and treatment of diseases related to immune maturation and  
CC regulation of immune response towards self and nonself. They can be used  
CC particularly in the diagnosis and treatment of autoimmune  
CC polyendocrinopathy candidiasis ectodermal dysplasia (APECED) (also known  
CC as autoimmune polyglandular syndrome type I (APS I))  
XX  
SQ Sequence 2036 BP; 359 A; 718 C; 649 G; 310 T; 0 U; 0 Other:  
Query Match 100.0%; Score 2036; DB 2; Length 2036;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2036; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGACCGGGAGACGGCGGGCGGACAGCCGGCCGGAGCCCAAGCCCGCGGAGCC 60  
DB 1 AGACCGGGAGAGCGGGGCTGGCCAGTGTCCGGGAGCCACCGCTCGGCCCAAGCCCGGAGCC 60  
QY 61 CGAGGCGCAAGCGAGGGGCTGCCAGTGTCCCGGGAGCCACCGGCTCGGCCCAAGCCCGGAGCC 120  
DB 61 CGAGGCGCAAGCGAGGGGCTGGCCAGTGTCCCGGGAGCCACCGCTCGGCCCAAGCCCGGAGCC 120  
QY 121 TCCCGCGGCCACCCCGATGGCGAGCGAGCGGCGCTTACGCGGCTTCTGAGGCTGGACCG 180  
DB 121 TCCCGCGGCCACCCCGATGGCGAGCGAGCGGCGCTTACGCGGCTTCTGAGGCTGGACCG 180  
QY 181 CAGGAGATCGCGGTGGCGGTGGAGCAGCGCTTCCACTGCTGCAAGCGGTGGCTTACCA 240  
DB 181 CAGGAGATCGCGGTGGCGGTGGAGCAGCGCTTCCACTGCTGCAAGCGGTGGCTTACCA 240  
QY 241 CGAGCGGTGCTCCCGAGGAGCAAGTTTCAAGAGAGCGCTTCACTGAAGAAAGAGAGGCGT 300  
DB 241 CGAGCGGTGCTCCCGAGGAGCAAGTTTCAAGAGAGCGCTTCACTGAAGAAAGAGAGGCGT 300  
QY 301 CCCCCAGGCTTCCACGCGCTCTGTCTGGCTGTGACCCAGAGATCCCAAGCCATCTT 360  
DB 301 CCCCCAGGCTTCCACGCGCTCTGTCTGGCTGTGACCCAGAGATCCCAAGCCATCTT 360  
QY 361 GGAATTCTGGAGGGTCTGTTCAGAGACTTAACTTGAAGCGCTTATGGCGGCTGCAAGCC 420  
DB 361 GGAATTCTGGAGGGTCTGTTCAGAGACTTAACTTGAAGCGCTTATGGCGGCTGCAAGCC 420  
QY 421 CATCTGAGCAGCTTCCCAAGATGTGACCTCAGCCAGCCCGGAAAGGAGGAGAGCC 480  
DB 421 CATCTGAGCAGCTTCCCAAGATGTGACCTCAGCCAGCCCGGAAAGGAGGAGAGCC 480  
QY 481 CCGGCGCGTCCCAAGAGGCTTGTGTAACCGCAACCAAGCTCCCAAGAGAGAGGCTT 540  
DB 481 CCGGCGCGTCCCAAGAGGCTTGTGTAACCGCAACCAAGCTCCCAAGAGAGAGGCTT 540  
QY 541 AGAAGAGGCTCGAGCTGCGCGCCAGCAGGCTGACTCCAAAGGGGACCGCCAGCCAG 600  
DB 541 AGAAGAGGCTCGAGCTGCGCGCCAGCAGGCTGACTCCAAAGGGGACCGCCAGCCAG 600  
QY 601 CTCTCAACTGAAGGCGCAAGCCCGCAAGAGCGGAGAGAGGAGGAGAGGAGGCT 660  
DB 601 CTCTCAACTGAAGGCGCAAGCCCGCAAGAGCGGAGAGAGGAGGAGAGGAGGCT 660  
QY 661 TTCACCTCGGGAACGGGATTCAAGCAATGTCAAGCTTCAAGTCAAGAGAGCTGTGCAATGTC 720  
DB 661 TTCACCTCGGGAACGGGATTCAAGCAATGTCAAGCTTCAAGTCAAGAGAGCTGTGCAATGTC 720  
QY 721 CTCGGGAGAGCTCCCGGAGAGCCGAGAGGCGCTGGAGAGGAGATCTTCAATCCAGAGAGTGT 780  
DB 721 CTCGGGAGAGCTCCCGGAGAGCCGAGAGGCGCTGGAGAGGAGATCTTCAATCCAGAGAGTGT 780  
QY 781 TGAGTCAGGCGGCTCCAGAAAGTGCATCCAGTTGGCGGGAGTTCATCCTCCAGCAA 840

DB 781 TGAGTCAGGCGGCTCCAGAAAGTGCATCCAGTTGGCGGGAGTTCATCCTCCAGCAA 840  
QY 841 GTTCGAAGACTCCGAGAGTGGGAAAGCAAGGCCCGCAGCAGTGGCCCAAGGCTCT 900  
DB 841 GTTCGAAGACTCCGAGAGTGGGAAAGCAAGGCCCGCAGCAGTGGCCCAAGGCTCT 900  
QY 901 GGTTCAGGCAAGGAGAGCCAGAGGCGTGGCCCCCGGTGAGAGTGAAGCTTGAAGCTGGCCA 960  
DB 901 GGTTCAGGCAAGGAGAGCCAGAGGCGTGGCCCCCGGTGAGAGTGAAGCTTGAAGCTGGCCA 960  
QY 961 GCAGGAGCAGGTTCCGCGCCCTCTGAGCCCTCCCAAGTGCAGCCCAAGCTTCAAGAAAGAA 1020  
DB 961 GCAGGAGCAGGTTCCGCGCCCTCTGAGCCCTCCCAAGTGCAGCCCAAGCTTCAAGAAAGAA 1020  
QY 1021 TGAGGACGAGTGTGCGGTGTGTGCGGAGCGGAGAGCTCATCTGTGACGAGTGTGCC 1080  
DB 1021 TGAGGACGAGTGTGCGGTGTGTGCGGAGCGGAGAGCTCATCTGTGACGAGTGTGCC 1080  
QY 1081 TGGGCGCTTCCACTGCGCTGTGCTGCTCCCTCCGCTCCGAGAGATCCCAAGTGGAGCTTG 1140  
DB 1081 TGGGCGCTTCCACTGCGCTGTGCTGCTCCCTCCGCTCCGAGAGATCCCAAGTGGAGCTTG 1140  
QY 1141 GAGTGTCTCAGTGTGCTGCAAGCAAGTCCAGAGAGTGCAGCCCGGAGAGAGAGCC 1200  
DB 1141 GAGTGTCTCAGTGTGCTGCAAGCAAGTCCAGAGAGTGCAGCCCGGAGAGAGAGAGCC 1200  
QY 1201 CCGGCGCCAGAGAGCAAGCCGTTGAGAGACCCCGCTCCCGGGGCTTAAAGTGGCGGAGAG 1260  
DB 1201 CCGGCGCCAGAGAGCAAGCCGTTGAGAGACCCCGCTCCCGGGGCTTAAAGTGGCGGAGAG 1260  
QY 1261 GAGAGTGAAGGTTCACCTGGGAGAACCCCTACCGGAGTGAACAAGACTTGTCTACAA 1320  
DB 1261 GAGAGTGAAGGTTCACCTGGGAGAACCCCTACCGGAGTGAACAAGACTTGTCTACAA 1320  
QY 1321 GCACCTGCGGCTCCGCTTCTGCAAGCCCGCTGCAAGGAGTGAAGTCTCCGAGCCCTGCA 1380  
DB 1321 GCACCTGCGGCTCCGCTTCTGCAAGCCCGCTGCAAGGAGTGAAGTCTCCGAGCCCTGCA 1380  
QY 1381 CCCCCACTGTGTGTGGGTCTGAGAGGTCAAGAGAACTGTGCTGTGCGGCTTGGCGG 1440  
DB 1381 CCCCCACTGTGTGTGGGTCTGAGAGGTCAAGAGAACTGTGCTGTGCGGCTTGGCGG 1440  
QY 1441 GGTGTGCGGAGATGTGTAACGAGAGTGTGAGGTGTACTACCTGCGCGCTTTCACATG 1500  
DB 1441 GGTGTGCGGAGATGTGTAACGAGAGTGTGAGGTGTACTACCTGCGCGCTTTCACATG 1500  
QY 1501 GCGCTGCCACTTCCAGCGGAGACTCCCGGAGCGGAGACGGGCTGTGCGTGCAGATCTTG 1560  
DB 1501 GCGCTGCCACTTCCAGCGGAGACTCCCGGAGCGGAGACGGGCTGTGCGTGCAGATCTTG 1560  
QY 1561 CTTCAGAGAGCTGACCCAGCCCTGTGTGAGAGGAGGTGTGAGCCCAAGCCCGGCT 1620  
DB 1561 CTTCAGAGAGCTGACCCAGCCCTGTGTGAGAGGAGGTGTGAGCCCAAGCCCGGCT 1620  
QY 1621 GGCCTCTGGGCTTGCAGAGAGTGAACATGCACTGCAAGCCCGCTTTCACAGAGGATGA 1680  
DB 1621 GGCCTCTGGGCTTGCAGAGAGTGAACATGCACTGCAAGCCCGCTTTCACAGAGGATGA 1680  
QY 1681 CCTGAGTCCCTTTCAGAGAGCAACCTTCATGAGCAATCTGAGTGGGCAATCAAGAG 1740  
DB 1681 CCTGAGTCCCTTTCAGAGAGCAACCTTCATGAGCAATCTGAGTGGGCAATCAAGAG 1740  
QY 1741 CATGCGCGCTCCGCGGCGCCCTTCCCTCTGACCCCAAGATGAGCGGAGCATGACGCTC 1800  
DB 1741 CATGCGCGCTCCGCGGCGCCCTTCCCTCTGACCCCAAGATGAGCGGAGCATGACGCTC 1800  
QY 1801 TGATGAGAGAGTCTGAGAGAGCAACTCTTCTCAAGTCTGAGAGCGCGCGGCTGGG 1860  
DB 1801 TGATGAGAGAGTCTGAGAGAGCAACTCTTCTCAAGTCTGAGAGCGCGCGGCTGGG 1860  
QY 1861 ATCAAGAGAGGAGCAGCGCACTCTGTGCAAGTGTGCGGCTGTAAACAGCTTGTGTTTC 1920

```

Db      1861 ATCAGAGAGGAGCAGCGCCCTTGTGTCAGTGTGCGCTGTAACAGCTGTGTTC 1920
QY      1921 TGGGGACACGACATCATGTGCTGGAATTAACCTGCCACTTCTCTACTGGA 1980
Db      1921 TGGGGACACGACATCATGTGCTGGAATTAACCTGCCACTTCTCTACTGGA 1980
QY      1981 AGTCCCGGAGGAGCTCTCTGCTGTGACACTACTATAAATTAATAATTAGCTG 2036
Db      1981 AGTCCCGGAGGAGCTCTCTGCTGTGACACTACTATAAATTAATAATTAGCTG 2036

RESULT 2
AA58605
ID      AA58605 standard; cDNA; 2245 BP.
XX
AC      AA58605;
XX
XX      16-AUG-1999 (first entry)
XX
DE      Human autoimmune polyglutandular disease type 1 (APGD1) cDNA.
XX
XX
XX      Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy; APECED;
KW      autoimmune polyglutandular disease type 1; APGD1; AIRE; human;
KW      transcription factor; autoimmune disease; diagnosis; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PH      Key
FT      CDS
FT      121..1758
FT      /tag= a
FT      /product= "human_APGD1"
FT      /replace(1358,T)
FT      /tag= k
FT      /replace(1374,G)
FT      /tag= l
FT      /replace(1708,T)
FT      /tag= o
FT      /note= "silent polymorphism"
FT      /replace(801,T)
FT      /tag= p
FT      /note= "silent polymorphism"
FT      /replace(889,T)
FT      /tag= j
FT      /note= "alters Arg codon to STOP, and truncated 256 aa
FT      protein"
FT      1051
FT      /tag= g
FT      /note= "1-nucleotide deletion"
FT      /replace(1052,A)
FT      /tag= m
FT      1085..1097
FT      /tag= f
FT      /note= "13-nucleotide deletion, results in frameshift and
FT      truncated 372 aa protein"
FT      /tag= b
FT      /note= "4-nucleotide insertion, results in frameshift and
FT      truncated 371 aa protein"
FT      /replace(1090,CTTG)
FT      /tag= c
FT      /note= "4-nucleotide insertion, results in frameshift and
FT      truncated protein"
FT      /replace(1094,A)
FT      /tag= n
FT      /replace(1284,A)
FT      /tag= d
FT      /note= "1-nucleotide insertion, results in frameshift and
FT      truncated 422 aa protein"
FT      1309
FT      /tag= h
FT      /note= "1-nucleotide deletion"
FT      1313
FT      /tag= i

```

```

FT      /note= "1-nucleotide deletion, results in frameshift and
FT      truncated 478 aa protein"
FT      /replace(1317,T)
FT      /tag= q
FT      /note= "silent polymorphism"
FT      /replace(1365,C)
FT      /tag= e
FT      /note= "1-nucleotide insertion"
FT      /replace(1698,T)
FT      /tag= r
FT      /note= "silent polymorphism"
FT      1995
FT      /tag= s
FT      /note= "Alu sequence overlapping with PKLX promoter
FT      starts at this position"
FT      2191..2195
FT      /tag= t

polyA_signal
FT
FT      MO9918197-A2.
XX
XX      15-APR-1999.
XX
XX      02-OCT-1998; 98WO-EP006294.
XX
XX      02-OCT-1997; 97EP-00117154.
XX      08-OCT-1997; 97EP-00117398.
XX      12-NOV-1997; 97EP-00119810.
XX
XX      (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX      (NAPU-) NAT PUBLIC HEALTH INST.
XX
XX      Peltonen L, Aaltonen J, Björnses P, Perheentupa J, Palotie A,
XX      Horelli-Kuitunen N, Yaspo M, Lehrach H;
XX      WPI; 1999-287735/24.
XX      P-PSDB; AAY06073.
XX
XX      New polypeptide which co-segregates in mutated form.
XX
XX      Claim 1, Fig 2A; 77pp; English.
XX
XX      This is the nucleotide sequence of human autoimmune polyglutandular disease
XX      type 1 (APGD1 or AIRE) cDNA. It is a composite of 2 clones, which were
XX      obtained from a human adult thymus cDNA library by PCR (see AA58609-10),
XX      and of a 3'UTR extension PCR product. The encoded protein (see AAY06073)
XX      is a transcription factor or transcription-associated factor that may
XX      associate with vimentin fibres, perhaps as part of a docking mechanism
XX      regulating nuclear translocation. Aggregates of the mutated protein may
XX      prevent formation of vimentin intermediate filaments. Mutated APGD1
XX      polypeptides co-segregate with autoimmune polyendocrinopathy candidiasis
XX      ectodermal dystrophy (APECED). The invention provides vectors and host
XX      cells for preparation of APGD1 polypeptides. APECED-associated mutations
XX      found in the APGD1 gene include: an insertion of 4 nucleotides (CTTG),
XX      normally found at position 1086-1089, at nucleotide position 1085 or 1090
XX      ; an insertion of an adenosine at position 1284; an insertion of a
XX      cytosine at position 1365; a deletion of nucleotides 1085-1097; a
XX      deletion of the thymidine at position 1051; a deletion of the cytosine at
XX      position 1309 or 1313; a cytosine to thymidine exchange at position 889;
XX      a guanosine to thymidine exchange at position 358; an adenosine to
XX      cytosine exchange at position 374; a guanosine to adenosine exchange at
XX      position 1052; and a cytosine to adenosine exchange at position 1094. A
XX      claimed method for testing for carriership of APECED or for the disease
XX      state involves testing for a mutation in the APGD1 gene or for a mutated
XX      form of the APGD1 polypeptide. APGD1 nucleic acid can be used in gene
XX      therapy
XX
SQ      Sequence 2245 BP; 419 A; 765 C; 697 G; 364 T; 0 U; 0 Other;
XX
Query Match      96.7%; Score 1969; DB 2; Length 2245;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      17 CGAGGACACGAGCGGAGAGCCCAAGCCCGCGGAGACCGAGGCCAAGGAGG 76

```

Db 1 CGGGCCGACAGCCGGGCGGAGGCCCAAGCCCGCGGAGCCCAAGCCAGAGG 60  
QY 77 GCTGCAGTGTCCCGGAGCCCAACCGGCTCGCCCAAGCCCGGGGTCCCGGCCCAACCC 136  
Db 61 GCTGCCAGTGTCCCGGAGCCCAACCGGCTCGCCCAAGCCCGGGGTCCCGGCCCAACCC 120  
QY 137 ATGGCGACGAGCCGGGCGCTACGCGCGCTTCTAGAGCTGCAACGCAACGAGATCGCGTG 196  
Db 121 ATGGCGACGAGCCGGGCGCTACGCGCGCTTCTAGAGCTGCAACGCAACGAGATCGCGTG 180  
QY 197 GCCGTGACAGGCGCTTCCCACTGTGACAGCGGCTGGCTGACACGACGAGTCCCGAG 256  
Db 181 GCCGTGACAGGCGCTTCCCACTGTGACAGCGGCTGGCTGACACGAGTGGTCCCGAG 240  
QY 257 GACAACTTTCAAGAGAAGCTTCACTGAAGAAAAGAGGCTGCCCAAGGCTTTCAC 316  
Db 241 GACAACTTTCAAGAGAAGCTTCACTGAAGAAAAGAGGCTGCCCAAGGCTTTCAC 300  
QY 317 GCCCTCTGTCTGTGCTGCTGACCCAGACTCCACAGCCATCTTGAATCTTGGAGGGTG 376  
Db 301 GCCCTCTGTCTGTGCTGCTGACCCAGACTCCACAGCCATCTTGAATCTTGGAGGGTG 360  
QY 377 CTGTTCAAGAGACTCAACCTGAGCGCTATGCGCGCTGAGCCCATCTTGAAGAGCTTC 436  
Db 361 CTGTTCAAGAGACTCAACCTGAGCGCTATGCGCGCTGAGCCCATCTTGAAGAGCTTC 420  
QY 437 CCCAAAGATGTGACTCAGCCAGCCCGGAGAGGAGAGAGCCCGCGCTGCCCAAG 496  
Db 421 CCCAAAGATGTGACTCAGCCAGCCCGGAGAGGAGAGAGCCCGCGCTGCCCAAG 480  
QY 497 GCTTTGGTACCCGACCCAGACTTCCCAACAGAGAGAGGCTCAGAGAGGCTCGAGCT 556  
Db 481 GCTTTGGTACCCGACCCAGACTTCCCAACAGAGAGAGGCTCAGAGAGGCTCGAGCT 540  
QY 557 GCCGCGCAGAGCCCTGACTCCAAAGGGGACCGGACCGAGGCTCTCAACTGAAGGCG 616  
Db 541 GCCGCGCAGAGCCCTGACTCCAAAGGGGACCGGACCGAGGCTCTCAACTGAAGGCG 600  
QY 617 AAGCCCCCAAGAGAGCCGAGAGAGCGAGAGAGAGAGAGGCTTCACTCGGGAACGG 676  
Db 601 AAGCCCCCAAGAGAGCCGAGAGAGCGAGAGAGAGAGAGGCTTCACTCGGGAACGG 660  
QY 677 ATTCAAGACATGTCACTTCACTGACAGAGAGCTGTGGCCATGTCTTCCGGGAGCTCCG 736  
Db 661 ATTCAAGACATGTCACTTCACTGACAGAGAGCTGTGGCCATGTCTTCCGGGAGCTCCG 720  
QY 737 GGAGCCGAGGGGCGGTGAGGGGATCTCATTCAGACAGGTGTTGAGTCAAGGCGGCTCC 796  
Db 721 GGAGCCGAGGGGCGGTGAGGGGATCTCATTCAGACAGGTGTTGAGTCAAGGCGGCTCC 780  
QY 797 AAGAAATGATCCAGGTTGGCGGAGATTCTACATCTCCAGCAAGTTCCAGAGATCCCGG 856  
Db 781 AAGAAATGATCCAGGTTGGCGGAGATTCTACATCTCCAGCAAGTTCCAGAGATCCCGG 840  
QY 857 AGTGGGAAGAACAGAGCCCGGACAGAGAGTGGCCGAGGCTTGTGTTGAGGCAAGGA 916  
Db 841 AGTGGGAAGAACAGAGCCCGGACAGAGAGTGGCCGAGGCTTGTGTTGAGGCAAGGA 900  
QY 917 GCCCAGAGGCGGTGCCCCCGGTGAGGTTGAGGCTGAGGCTGAGAGGAGGAGGCTTCC 976  
Db 901 GCCCAGAGGCGGTGCCCCCGGTGAGGTTGAGGCTGAGGCTGAGAGGAGGAGGCTTCC 960  
QY 977 GCCCCTCTGAGCCCTCCCAAGTACCCCAAGCTTCAACAGAGATGAGAGCGAGTGTGC 1036  
Db 961 GCCCCTCTGAGCCCTCCCAAGTACCCCAAGCTTCAACAGAGATGAGAGCGAGTGTGC 1020  
QY 1037 GTGTGTGGGAGCGGCGGAGACTCATCTGTGTGACGGCTGCCCTTGGGCTTCCACTG 1096  
Db 1021 GTGTGTGGGAGCGGCGGAGACTCATCTGTGTGACGGCTGCCCTTGGGCTTCCACTG 1080  
QY 1097 GCTGCTGTCCCTCGGCTCGGAGAGATCCCAAGTGGGAGCTGAGAGTGTCTCAAGTGC 1156

Db 1081 GCTGCTGTCCCTCCGCTCCGGAGATCCCAAGTGGGACCTGAGAGTGTCTCAAGTGC 1140  
QY 1157 CTGACAGCAACAGTCCAGAGAGTGACGCCCGGAGAGAGAGAGAGAGAGAGAGCA 1216  
Db 1141 CTGACAGCAACAGTCCAGAGAGTGACGCCCGGAGAGAGAGAGAGAGAGAGAGCA 1200  
QY 1217 CCCGTGAGAACCCCGTCCCGGCGGTAGTGTGGCGGAGAGAGAGTAAAGAGTCCA 1276  
Db 1201 CCCGTGAGAACCCCGTCCCGGCGGTAGTGTGGCGGAGAGAGAGTAAAGAGTCCA 1260  
QY 1277 CCTGGGAAACCCCTAGCCGAGCATGACAGACTTGTCTAAGACACTGCGGCTCCG 1336  
Db 1261 CCTGGGAAACCCCTAGCCGAGCATGACAGACTTGTCTAAGACACTGCGGCTCCG 1320  
QY 1337 CCTTTCAGAGCCCGCTGACAGGCTGAGACTCTTGGCCCTGACCCCTTACTGTGTG 1396  
Db 1321 CCTTTCAGAGCCCGCTGACAGGCTGAGACTCTTGGCCCTGACCCCTTACTGTGTG 1380  
QY 1397 GGTCTGAGAGGTACAGAGAACCTGGCTCCTGGTGGCGGTTGGGGGTGGCGAGATGGT 1456  
Db 1381 GGTCTGAGAGGTACAGAGAACCTGGCTCCTGGTGGCGGTTGGGGGTGGCGAGATGGT 1440  
QY 1457 ACGAGAGTGTGCTGAGTACTGACTGCGCGGCTGCTTCACTGAGGCTGCCACTTCCA 1516  
Db 1441 ACGAGAGTGTGCTGAGTACTGACTGCGCGGCTGCTTCACTGAGGCTGCCACTTCCA 1500  
QY 1517 GCCGCACTTCCCGGCGGAGAGCGGCTGCGTGCAGATCTCTGCTCAGAGAGAGTACC 1576  
Db 1501 GCCGCACTTCCCGGCGGAGAGCGGCTGCGTGCAGATCTCTGCTCAGAGAGAGTACC 1560  
QY 1577 CGAGCCCTGTGAGAGGGGTGTGAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1636  
Db 1561 CGAGCCCTGTGAGAGGGGTGTGAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1637 AAGATGACACTGACAGTACAGAGCCGCTCTGCAACAGGAGTGAACCTGAGTCCCTTG 1696  
Db 1621 AAGATGACACTGACAGTACAGAGCCGCTCTGCAACAGGAGTGAACCTGAGTCCCTTG 1680  
QY 1697 ACGAGACACCTTGTGATGAGCATCTGCAAGTGGGCAATCAGAGCATGGCCGCTCCGCG 1756  
Db 1681 ACGAGACACCTTGTGATGAGCATCTGCAAGTGGGCAATCAGAGCATGGCCGCTCCGCG 1740  
QY 1757 GCCCCTTCCCTCCGAGCCCAAGATGGCGGAGATGAGAGTCTGATGAGAGAGTGC 1816  
Db 1741 GCCCCTTCCCTCCGAGCCCAAGATGGCGGAGATGAGAGTCTGATGAGAGAGTGC 1800  
QY 1817 AGAAGGACACTCTTCTCACTGAGTCTGAGAGCGGCGGCTGGGATCAAGAGGAGAG 1876  
Db 1801 AGAAGGACACTCTTCTCACTGAGTCTGAGAGCGGCGGCTGGGATCAAGAGGAGAG 1860  
QY 1877 GGCACCTTGTGAGTGTGAGTGTGAGTGTGAAACAGCTCTGTGTTCTGGGAGACAGCCAT 1936  
Db 1861 GGCACCTTGTGAGTGTGAGTGTGAGTGTGAAACAGCTCTGTGTTCTGGGAGACAGCCAT 1920  
QY 1937 CATGTGCTGTGAAATTAACCCCTGCCCCCACTTCTCACTGAGAGTCCCGGGAGCCCTC 1996  
Db 1921 CATGTGCTGTGAAATTAACCCCTGCCCCCACTTCTCACTGAGAGTCCCGGGAGCCCTC 1980  
QY 1997 TCCCTGCTGTGACTACTAATAAATATTAATTAAGCTG 2036  
Db 1981 TCCCTGCTGTGACTACTAATAAATATTAATTAAGCTG 2020

RESULT 3  
AAK26937  
ID AAK26937 standard; DNA; 1545 BP.  
XX  
XX AAK26937;  
XX  
AC 24-JUN-1999 (first entry)  
XX  
DT  
XX  
DE CDNA encoding a human autoimmune regulator-2 (AIR-2) protein.  
XX

KM Autoimmune regulator-2; AIR-2; immune maturation; immune response;  
KM disease; autoimmune polyendocrinopathy candidiasis ectodermal dystrophy;  
KM APECED; autoimmune polyglandular syndrome type I; APS I; ss.

XX Homo sapiens.

OS  
FH Key Location/Qualifiers  
FT CDS 237..1283  
FT /tag= a "AIR-2"  
FT /product= "AIR-2"

PN WO915559-A1.

PD 01-APR-1999.

PF 23-SEP-1998; 98WO-FI000749.

PR 23-SEP-1997; 97FI-00003762.

PA (FIIIM-) FINNISH IMMUNOTECHNOLOGY LTD.

XX Krohn K, Heino M, Peterson P, Scott H, Antonarakis S, Lalioti M;  
PI Shimizu N, Kudoh J;

DR MPI; 1999-244390/20.

DR P-PDB; AAY01713.

PT Autoimmune regulator 1 (AIR1) DNA sequence.

XX Claim 4; Page 27-29; 59pp; English.

CC The present sequence encodes an autoimmune regulator-2 (AIR-2) protein.  
CC The AIR polypeptides and polynucleotides can be used in methods for the  
CC diagnosis and treatment of diseases related to immune maturation and  
CC regulation of immune response towards self and nonself. They can be used  
CC particularly in the diagnosis and treatment of autoimmune  
CC polyendocrinopathy candidiasis ectodermal dystrophy (APECED) (also known  
CC as autoimmune polyglandular syndrome type I (APS I))

XX Sequence 1545 BP; 261 A; 485 C; 508 G; 291 T; 0 U; 0 Other;

Query Match 50.3%; Score 1025; DB 2; Length 1545;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1012 CCAGAGATGATGAGAGAGTGGCGGTGTGCGGACGGCGGAGACTCATCTGCTGTA 1071

DB 521 CCAGAGAAATGAGAGAGAGTGTGCTGTGCGGACGGCGGAGACTCATCTGCTGTA 580

QY 1072 CGGCTCCCTCGGCGCTTCACCTGAGCTGTGCTCCCTCCGCTCCGGGAGATCCCGAG 1131

DB 581 CGGCTCCCTCGGCGCTTCACCTGAGCTGTGCTCCCTCCGCTCCGGGAGATCCCGAG 640

QY 1132 TGGGACTGAGAGTGTCTCCAGCTGTGTCAGAGCAAGTCCAGAGAGTGCAGCCCGAGC 1191

DB 641 TGGGACTGAGAGTGTCTCCAGCTGTGTCAGAGCAAGTCCAGAGAGTGCAGCCCGAGC 700

QY 1192 AAGAGAGCCCGCGCGCGCGAGGACCCCTGAGAGACCCCGCTCCCGCGGAGTTAGGTC 1251

DB 701 AAGAGAGCCCGCGCGCGCGAGGACCCCTGAGAGACCCCGCTCCCGCGGAGTTAGGTC 760

QY 1252 GCGGAGAGAGAGTGAAGAGTCCACTGAGGAAACCTTAGCGGATGAGACGACTCTC 1311

DB 761 GCGGAGAGAGAGTGAAGAGTCCACTGAGGAAACCTTAGCGGATGAGACGACTCTC 820

QY 1312 TGTCTACAAAGCACTGCGCGCTCCGCTTCTGAGGCCCGCTGTCAGAGGCTGACTCTC 1371

DB 821 TGTCTACAAAGCACTGCGCGCTCCGCTTCTGAGGCCCGCTGTCAGAGGCTGACTCTC 880

QY 1372 GGCCTTGACCCCTGACTGTGTGTGGTCTTGAAGGTGACGAAACCTGGCTCTGTGTC 1431

DB 881 GGCCTTGACCCCTGACTGTGTGTGGTCTTGAAGGTGACGAAACCTGGCTCTGTGTC 940

QY 1432 GCGTTCGGGGGTGTGTCGGAGATGTAGCAGAGCTGTGCGGTGTACTCATCTGCGCGCTGC 1491

DB 941 GCGTTCGGGGGTGTGTCGGAGATGTAGCAGAGCTGTGCGGTGTACTCATCTGCGCGCTGC 1000

QY 1492 CTTTCACATGCGCTGTCCACTTCCAGCCGGACCTTCCCGGCCCGGAGCGGCTCGCTG 1551

DB 1001 CTTTCACATGCGCTGTCCACTTCCAGCCGGACCTTCCCGGCCCGGAGCGGCTCGCTG 1060

QY 1552 CAGATCTGCTCAGAGAGACGTGACCCCAAGCCCCCTGTGAGAGGGGGGTCGGCCCCAGACC 1611

DB 1061 CAGATCTGCTCAGAGAGACGTGACCCCAAGCCCCCTGTGAGAGGGGGGTCGGCCCCAGACC 1120

QY 1612 CCGCCGCTGCGCCCTGCGGCTGCCAAGATGACCTGCACTGACGAGCCGCTGTGA 1671

DB 1121 CCGCCGCTGCGCCCTGCGGCTGCCAAGATGACCTGCACTGACGAGCCGCTGTGA 1180

QY 1672 CAGGATGACCTGAGATCCCTTCTGAGCGAGACACCTTTCATGATGCACTGAGTGGC 1731

DB 1181 CAGGATGACCTGAGATCCCTTCTGAGCGAGACACCTTTCATGATGCACTGAGTGGC 1240

QY 1732 CATCCAGAGCATGCGCCGTCGGCGGCCCCCTTCCCTCCCTGACCCCAATGCGCGGAC 1791

DB 1241 CATCCAGAGCATGCGCCGTCGGCGGCCCCCTTCCCTCCCTGACCCCAATGCGCGGAC 1300

QY 1792 ATGCACTCTGTAGAGAGATGCTGAGAAAGACACTTCTCTCACTGCTGGAAGCCGG 1851

DB 1301 ATGCACTCTGTAGAGAGATGCTGAGAAAGACACTTCTCTCACTGCTGGAAGCCGG 1360

QY 1852 CCGGCTGGAGTCAAGAGAGGAGACAGCGCCACTCTTGTCAAGTGTCTGCGCTGTAACAGT 1911

DB 1361 CCGGCTGGAGTCAAGAGAGGAGACAGCGCCACTCTTGTCAAGTGTCTGCGCTGTAACAGT 1420

QY 1912 CTGTCTTTCTGGGGACACAGGCATCATGTGCTGGAATTAACCTGCCCACTTCTGC 1971

DB 1421 CTGTCTTTCTGGGGACACAGGCATCATGTGCTGGAATTAACCTGCCCACTTCTGC 1480

QY 1972 TACTGTGAAGTCCCGGAGGCTCTCTTGTGCTGTGACCTTAATAATAATAATT 2031

DB 1481 TACTGTGAAGTCCCGGAGGCTCTCTTGTGCTGTGACCTTAATAATAATAATAATT 1540

QY 2032 AGCTG 2036

DB 1541 AGCTG 1545

RESULT 4  
AAK26938  
ID AAK26938 standard; DNA; 1463 BP.

XX AAK26938;

XX 24-JUN-1999 (first entry)

XX cDNA encoding a human autoimmune regulator-3 (AIR-3) protein.

XX Autoimmune regulator-3; AIR-3; immune maturation; immune response;

XX disease; autoimmune polyendocrinopathy candidiasis ectodermal dystrophy;

XX APECED; autoimmune polyglandular syndrome type I; APS I; ss.

XX Homo sapiens.

OS  
FH Key Location/Qualifiers  
FT CDS 237..1001  
FT /tag= a "AIR-3"  
FT /product= "AIR-3"

PN WO915559-A1.

PD 01-APR-1999.

PF 23-SEP-1998; 98WO-FI000749.

PR 23-SEP-1997; 97FI-00003762.

XX (FIM-) FINNISH IMMUNOTECHNOLOGY LTD.  
 PA Krohn K, Heino M, Peterson P, Scott H, Antonarakis S, Lallioti M,  
 PI Shimizu N, Kudoh J;  
 XX WPI: 1999-244390/20.  
 DR P-PSDB; AAY01714.  
 XX Autoimmune regulator 1 (AIR1) DNA sequence.  
 PT Claim 4; Page 31-32; 59pp; English.  
 XX  
 PS The present sequence encodes an autoimmune regulator-3 (AIR-3) protein.  
 CC The AIR polypeptides and polynucleotides can be used in methods for the  
 CC diagnosis and treatment of diseases related to immune maturation and  
 CC regulation of immune response towards self and nonself. They can be used  
 CC particularly in the diagnosis and treatment of autoimmune  
 CC polyendocrinopathy candidiasis ectodermal dysplasia (APECED) (also known  
 CC as autoimmune polyglandular syndrome type I (APS I))  
 CC  
 XX Sequence 1463 BP; 245 A; 456 C; 486 G; 276 T; 0 U; 0 Other;  
 SQ  
 Query Match 34.0%; Score 692; DB 2; Length 1463;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1345 AGCCCCGCTCCAGAGGGCTGAGTCTCTGGCCCTGACACCCCTTACTGTGTGGTCTGA 1404  
 DB 772 ACCCCCGCTCCAGAGGGCTGAGTCTCTGGCCCTGACACCCCTTACTGTGTGGTCTGA 831  
 QY 1405 GGGTACGACGAACCTGGCTCTGTGTGGGCTTGGGGGTGTGGGAGATGCTACGACGT 1464  
 DB 832 GGGTACGACGAACCTGGCTCTGTGTGGGCTTGGGGGTGTGGGAGATGCTACGACGT 891  
 QY 1465 GCTGCGGTGTACTACTACTGCGCGCTCTTCCACTGGCGCTGCACTTCCAGCCGCGAC 1524  
 DB 892 GCTGCGGTGTACTACTACTGCGCGCTCTTCCACTGGCGCTGCACTTCCAGCCGCGAC 951  
 QY 1525 CTCCCGGCGCGGAGCGGGCTTGGCTGACAGATCTCTGACGAGACGTGACCCGACGCC 1584  
 DB 952 CTCCCGGCGCGGAGCGGGCTTGGCTGACAGATCTCTGACGAGACGTGACCCGACGCC 1011  
 QY 1585 TGTGAGGGGGGTGTGTGGCCCGCCCGCGCTGGGCTTGGGCTTGGCAAGATGA 1644  
 DB 1012 TGTGAGGGGGGTGTGTGGCCCGCCCGCGCTGGGCTTGGGCTTGGCAAGATGA 1071  
 QY 1645 CACTGCACTGACGAGCCGCGCTTGGCAAGGATGACAGTCCCTTCTGAGGAGCA 1704  
 DB 1072 CACTGCACTGACGAGCCGCGCTTGGCAAGGATGACAGTCCCTTCTGAGGAGCA 1131  
 QY 1705 CACTTTCAGTGGATCTCTGAGTGGGCTTCCAGAGCATGGCCGTCGCGGCCCTT 1764  
 DB 1132 CACTTTCAGTGGATCTCTGAGTGGGCTTCCAGAGCATGGCCGTCGCGGCCCTT 1191  
 QY 1765 CCCCTTCTAACCAGATGGCGGGGACATGCACTTGAATGAGAGTCTTGAAGAGAC 1824  
 DB 1192 CCCCTTCTAACCAGATGGCGGGGACATGCACTTGAATGAGAGTCTTGAAGAGAC 1251  
 QY 1825 ACCCTCTTCTGAGTCTTGAAGCCGCGGCTGGATCATGAAGAGGGGACAGGCCACT 1884  
 DB 1252 ACCCTCTTCTGAGTCTTGAAGCCGCGGCTGGATCATGAAGAGGGGACAGGCCACT 1311  
 QY 1885 CTGTGACGTCTGGCTGTAAACAGCTGTGTGTTTGTGGGACACAGGCATCATGTGCC 1944  
 DB 1312 CTGTGACGTCTGGCTGTAAACAGCTGTGTGTTTGTGGGACACAGGCATCATGTGCC 1371  
 QY 1945 TGGAAATTTAAACCTGGCCCACTTCTCTAATCTTGAAGAGTCCCGGAGAGCTTCTTGGC 2004  
 DB 1372 TGGAAATTTAAACCTGGCCCACTTCTCTAATCTTGAAGAGTCCCGGAGAGCTTCTTGGC 1431  
 QY 2005 TGGTACCTACTATAAATAATAAATAATTAAGCTG 2036

DB 1432 TGGTACCTACTATAAATAATAAATAATTAAGCTG 1463  
 RESULT 5  
 ACH67354  
 ID ACH67354 standard; DNA; 556 BP.  
 XX  
 AC ACH67354;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DB Human genome derived single exon probe #549.  
 XX  
 KW Human; probe; 58; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003194704-A1.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 03-APR-2002; 2002US-00029386.  
 XX  
 PR 03-APR-2002; 2002US-00029386.  
 XX  
 PA (PENN/) PENN S G.  
 PA (RANK/) RANK D R.  
 PA (HANZ/) HANZEL D K.  
 PI Penn SG, Rank DR, Hanzel DK;  
 XX  
 DR WPI: 2004-119264/12.  
 XX  
 PT New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX  
 PS Claim 15; SEQ ID NO 549; 80bp; English.  
 XX  
 CC The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridizes under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subscription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probe may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterizing  
 CC alternative splicing events, in detecting and characterizing gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC segdata.uspto.gov/sequence.html?docID=20030194704  
XX  
SQ Sequence 556 BP; 102 A; 195 C; 163 G; 96 T; 0 U; 0 Other;  
Query Match 8.7%; Score 178; DB 12; Length 556;  
Best Local Similarity 100.0%; Pred. No. 2.9e-73;  
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 266 CAGGAGACGCTTCATCTGAAGAAAAGAGGCTGCCCGCCCTTCACGCTTCCTG 325  
DB 256 CAGGAGACGCTTCATCTGAAGAAAAGAGGCTGCCCGCCCTTCACGCTTCCTG 315  
QY 326 TCTGGCTGTGACCCAGGACTCCACAGCATCTTGAGCTTCGAGGGGCTGTCAAG 385  
DB 316 TCTGGCTGTGACCCAGGACTCCACAGCATCTTGAGCTTCGAGGGGCTGTCAAG 375  
QY 386 GACTACACCTGGAGCGCTATGCGCGCTGCAGGCCCATCTTGAGACAGCTTCCCAAG 443  
DB 376 GACTACACCTGGAGCGCTATGCGCGCTGCAGGCCCATCTTGAGACAGCTTCCCAAG 433  
RESULT 6  
ACH81059  
ID ACH81059 standard; DNA; 175 BP.  
XX  
AC ACH81059;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human genome derived single exon probe #14254.  
XX  
KW Human; probe; ss; gene expression; single exon probe; microarray;  
KM alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
DR WPI; 2004-119264/12.  
XX  
PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
XX  
PS Claim 1; SEQ ID NO 14254; 80pp; English.  
XX  
CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 688 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of

CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC segdata.uspto.gov/sequence.html?docID=20030194704  
XX  
SQ Sequence 175 BP; 35 A; 60 C; 48 G; 32 T; 0 U; 0 Other;  
Query Match 8.6%; Score 175; DB 12; Length 175;  
Best Local Similarity 100.0%; Pred. No. 7.9e-72;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 269 GAGACGCTTCATCTGAAGAAAAGAGGCTGCCCGCCCTTCACGCTTCCTGTCC 328  
DB 1 GAGACGCTTCATCTGAAGAAAAGAGGCTGCCCGCCCTTCACGCTTCCTGTCC 60  
QY 329 TGGCTGTGACCCAGGACTCCACAGCATCTTGAGCTTCGAGGGCTGTTCAGAGAC 388  
DB 61 TGGCTGTGACCCAGGACTCCACAGCATCTTGAGCTTCGAGGGCTGTTCAGAGAC 120  
QY 389 TACACCTGGAGCGCTATGCGCGCTGCAGGCCCATCTTGAGACAGCTTCCCAAG 443  
DB 121 TACACCTGGAGCGCTATGCGCGCTGCAGGCCCATCTTGAGACAGCTTCCCAAG 175  
RESULT 7  
ACH68194  
ID ACH68194 standard; DNA; 530 BP.  
XX  
AC ACH68194;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human genome derived single exon probe #1389.  
XX  
KW Human; probe; ss; gene expression; single exon probe; microarray;  
KM alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
DR WPI; 2004-119264/12.  
XX

XX New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX  
 PS Claim 15; SEQ ID NO 1389; 80bp; English.  
 XX  
 CC The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridizes under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subcription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterizing  
 CC alternative splicing events, in detecting and characterizing gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docid=20030194704  
 XX  
 SQ Sequence 530 BP; 87 A; 187 C; 170 G; 86 T; 0 U; 0 Other;  
 XX  
 Query Match 7.4%; Score 150; DB 12; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-60;  
 Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 786 CAGGCGGCTCCAGAAAGTCATCCAGGTTGGCGGGAGTTCTACCTCCAGCAAGTTGC 845  
 DB 103 CAGGCGGCTCCAGAAAGTCATCCAGGTTGGCGGGAGTTCTACCTCCAGCAAGTTGC 162  
 QY 846 AAGACTCCGCGAGTGGAGAAACAAGCCCGCAGCAGCACTGCCCCGAGCCCTCTGTTTC 905  
 DB 163 AAGACTCCGCGAGTGGAGAAACAAGCCCGCAGCAGCACTGCCCCGAGCCCTCTGTTTC 222  
 QY 906 GAGCCAAAGGAGCCCGAGGCGCTGCCCCCG 935  
 DB 223 GAGCCAAAGGAGCCCGAGGCGCTGCCCCCG 252  
 RESULT 8  
 ACH81896  
 ID ACH81896 standard; DNA; 146 BP.  
 XX  
 AC ACH81896;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DE Human genome derived single exon probe #15091.  
 XX

KW Human; probe; ss; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.  
 OS Homo sapiens.  
 XX  
 PN US2003194704-A1.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 03-APR-2002; 2002US-00029386.  
 XX  
 PR 03-APR-2002; 2002US-00029386.  
 XX  
 PA (PENN/) PENN S G.  
 PA (RANK/) RANK D R.  
 PI (HANK/) HANK D K.  
 PI Penn SG, Rank DR, Hanzel DK;  
 XX  
 DR WPI; 2004-119264/12.  
 XX  
 PT New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 PS Claim 1; SEQ ID NO 15091; 80bp; English.  
 XX  
 CC The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridizes under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subcription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterizing  
 CC alternative splicing events, in detecting and characterizing gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docid=20030194704  
 XX  
 SQ Sequence 146 BP; 32 A; 46 C; 49 G; 19 T; 0 U; 0 Other;  
 XX  
 Query Match 7.2%; Score 146; DB 12; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-58;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 789 GCGGCTCCAGAAAGTCATCCAGGTTGGCGGGAGTTCTACCTCCAGCAAGTTGCAAG 848  
 DB 1 GCGGCTCCAGAAAGTCATCCAGGTTGGCGGGAGTTCTACCTCCAGCAAGTTGCAAG 60



```
QY      849 ACTCCGCGAGTGGGAGAAACAAGCCGCGAGCAGTGGCCCGAAGCCTTGTTGAG 908
      |||
      61 ACTCCGCGAGTGGGAGAAACAAGCCGCGAGCAGTGGCCCGAAGCCTTGTTGAG 120
Db
QY      909 CCAAGGAGCCGAGCGCGCTGCCCC 934
      |||
      121 CCAAGGAGCCGAGCGCGCTGCCCC 146
Db

RESULT 9
ABN38615
ID      ABN38615 standard; DNA; 60 BP.
XX
XX
AC      ABN38615;
XX
XX      15-JUL-2002 (first entry)
DT
XX
XX      Human spliced transcript detection oligonucleotide SEQ ID NO:11363.
DE
XX
XX      Human; mouse; rat; splice transcript; detection; RNA transcript;
KM      splice variant; transcriptome; oligonucleotide library; ss.
XX
XX      Homo sapiens.
OS
XX      WO200210449-A2.
PN
XX      07-FEB-2002.
PD
XX      20-JUL-2001; 2001WO-IB001903.
PF
XX      28-JUL-2000; 2000US-0221607P.
PR      02-MAY-2001; 2001US-0287724P.
XX
XX      (COMP-) COMPUGEN INC.
PA
XX      Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX      WPI; 2002-257383/30.
DR
XX
XX      New oligonucleotide libraries comprising oligonucleotides which
PT      selectively hybridize to mRNAs transcribed from a transcription unit of a
PT      genome, useful for detecting tissue-, pathology-, and developmental-
PT      specific genes.
XX
XX      Example 1; SEQ ID NO 11363; 47bp; English.
PS
XX
XX      The present invention describes oligonucleotide libraries for detecting
CC      messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC      )transcriptome comprises messenger RNAs transcribed from multiple
CC      transcription units that populate a genome. The library comprises several
CC      oligonucleotides, each capable of hybridising selectively to a set of
CC      messenger RNAs transcribed from a given transcription unit of the genome,
CC      which encodes one or more messenger RNA splice variants. The
CC      oligonucleotide libraries are useful for detecting mRNAs from a
CC      biological sample, in expression profiling studies, in qualitatively or
CC      quantitatively characterising the corresponding transcriptome, and in
CC      detecting RNA transcripts and splice variants of human or animal
CC      transcriptomes. The libraries may also be used as specialised mini
CC      libraries to detect transcripts of a sub-transcriptome under a particular
CC      biological or pathological state, and so allowing the detection of tissue
CC      - and pathology-specific genes such as those genes only expressed in
CC      specific tissue under a specific pathological condition; to detect
CC      developmental specific genes; and to detect RNA transcripts and splice
CC      variants of a transcriptome of a patient suffering from a particular
CC      disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC      rats, humans and mice, which are used in the exemplification of the
CC      present invention. N.B. The sequence data for this patent did not form
CC      part of the printed specification, but was obtained in electronic format
CC      directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX      Sequence 60 BP; 13 A; 15 C; 16 G; 16 T; 0 U; 0 Other;
```

```
Query Match      2.9%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3e-17;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1894 GCTCGGCTGTAAACAGCTCTGTGTTTCTGGGACACAGCCATCATGTGCTGGAATT 1953
      |||
      1 GCTCGGCTGTAAACAGCTCTGTGTTTCTGGGACACAGCCATCATGTGCTGGAATT 60
Db

RESULT 10
AAK58606
ID      AAK58606 standard; cDNA; 1656 BP.
XX
XX
AC      AAK58606;
XX
XX      16-AUG-1999 (first entry)
DT
XX
XX      Mouse autoimmune polyglanular disease type 1 (APGD1) cDNA.
DE
XX
XX      Autoimmune polyendocrinopathy candidiasis ectodermal dysrophy; APECED;
KM      autoimmune polyglanular disease type 1; APGD1; AIRE;
KM      transcription factor; autoimmune disease; diagnosis; gene therapy; mouse;
KM      animal model; ss.
XX
XX      Mus sp.
OS
XX      MO9918197-A2.
PN
XX      15-APR-1999.
PD
XX      02-OCT-1998; 98WO-EP006294.
PF
XX      02-OCT-1997; 97EP-00117154.
PR      08-OCT-1997; 97EP-00117398.
PR      12-NOV-1997; 97EP-00119810.
XX
XX      (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA      (NAPU-) NAT PUBLIC HEALTH INST.
XX
XX      Peltonen L, Aaltonen J, Bjoerres P, Perhentupa J, Palotie A;
PI      Horelli-Kuitunen N, Yaebo M, Lehrach H;
XX
XX      WPI; 1999-287735/24.
DR      P-PSDB; AAY06074.
XX
XX      New polypeptide which co-segregates in mutated form.
PT
XX
XX      Claim 6; Fig 14A-C; 77bp; English.
PS
XX
XX      This is the nucleotide sequence of murine AIRE cDNA. Murine AIRE is the
CC      homologue of the human AIRE (or autoimmune polyglanular disease type 1,
CC      APGD1) gene (see AAK58605). The overall nucleotide sequence identity
CC      between the mouse AIRE coding sequence and that of the human is 76.67%.
CC      The AIRE gene was localised to chromosome 10. Human mutated APGD1 co-
CC      segregates with autoimmune polyendocrinopathy candidiasis ectodermal
CC      dysrophy (APECED). The murine homologue may be used to develop an animal
CC      model for APECED, to examine the events that lead to the development of
CC      APECED and possibly to develop agents for preventing and/or treating this
CC      autoimmune disease
XX
XX      Sequence 1656 BP; 344 A; 532 C; 479 G; 301 T; 0 U; 0 Other;
```

```
Query Match      1.6%; Score 32; DB 2; Length 1656;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1484 GCCGCTGCCTTCACTGAGCGCTGCCACTTCCC 1515
      |||
      1348 GCCGCTGCCTTCACTGAGCGCTGCCACTTCCC 1379
Db
```

```
RESULT 11
AAK42015/c
```





PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251688P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-465566/50.  
XX  
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,  
PT treating neural, immune system, muscular, reproductive, pulmonary,  
PT cardiovascular, renal, proliferative disorders and cancerous diseases.  
XX  
PS Disclosure; SEQ ID NO 2141; 1180pp; English.  
XX  
CC The present invention relates to the isolation of novel human enzyme  
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
CC isomerases or ligases. The sequences of the invention are useful in the  
CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
CC disorders including hyperproliferative disorders (e.g. cancer),  
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.  
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic  
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),  
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders  
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and  
CC infectious disorders (e.g. influenza). The polynucleotides of the  
CC invention can also be used in gene therapy. AAS41685-AAS42192 represent  
CC DNA sequences encoding for the novel human enzyme polypeptides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 9742 BP; 2382 A; 2564 C; 2516 G; 2280 T; 0 U; 0 Other;  
Query Match 1.3%; Score 26; DB 4; Length 9742;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2011 CCTACTAAATAATTAATAATTAGCTG 2036  
DB 2692 CCTACTAAATAATTAATAATTAGCTG 2667  
RESULT 12  
ID AAK71122 standard; DNA; 95 BP.  
XX  
XX AAK71122;  
AC  
DT 06-NOV-2001 (first entry)  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25934.  
DE  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KM cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200157182-A2.  
XX  
XX

PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0218865P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227099P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0231415P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233633P.  
PR 14-SEP-2000; 2000US-0233634P.  
PR 14-SEP-2000; 2000US-0233635P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241212P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0241826P.  
 PR 08-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251897P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR MPI; 2001-483426/52.

XX XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
 XX  
 XX Disclosure; SEQ ID NO 25934; 3071bp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 SO Sequence 95 BP; 20 A; 28 C; 20 G; 27 T; 0 U; 0 Other;  
 Query Match 1.2%; Score 25; DB 4; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 0.41;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTTACTAAATATATAAATTACTG 2036  
 DB 38 CTTACTAAATATATAAATTACTG 14  
 ID ABA16553/c  
 ID ABA16553 standard; DNA; 113 BP.  
 AC ABA16553;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Human nervous system related polynucleotide SEQ ID NO 8884.  
 XX  
 XX Human; noctropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;  
 KW antiparasitic; antitumor; antidiabetic; antitubercular; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antitubercular; anticonvulsant; antifungal;  
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
 OS Homo sapiens.  
 PN WO200159063-A2.  
 XX  
 PD 16-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001334.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214866P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.



CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anaemia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [http://wipo.int/patdb/published\\_pct\\_sequences](http://wipo.int/patdb/published_pct_sequences)  
XX  
SQ Sequence 113 BP; 22 A; 38 C; 22 G; 31 T; 0 U; 0 Other;  
  
Query Match 1.2%; Score 25; DB 5; Length 113;  
Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 2012 CTAAGTAAATATATATATATAGCTG 2036  
Db 38 CTAAGTAAATATATATATATAGCTG 14  
  
RESULT 14  
AAK74219  
ID AAK74219 standard; DNA; 147 BP.  
XX  
AC AAK74219; {  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29031.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229433P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236802P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239933P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.



PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234224P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236328P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250316P.  
PR 01-DEC-2000; 2000US-0250319P.  
PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
DR  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
PS Disclosure; SEQ ID NO 29030; 3071pp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 147 BP; 41 A; 33 C; 42 G; 31 T; 0 U; 0 Other;  
Query Match 1.2%; Score 25; DB 4; Length 147;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATAAATTAGCTG 2036  
Db 60 CTACTAAAAATATAAATTAGCTG 84  
RESULT 16  
AAAC12190  
ID AAC12190 standard; cDNA; 172 BP.  
XX  
AC AAC12190;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 16265.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.  
PA Dumas Milne Edwards J, Duclet A, Giordano J;  
PI MPI; 2000-500361/45.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
XX Claim 1; SEQ ID NO 16265; 71bp + Sequence Listing; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or poly(A<sup>+</sup> RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors  
XX  
SQ Sequence 172 BP; 50 A; 34 C; 43 G; 44 T; 0 U; 1 Other;  
  
Query Match 1.2%; Score 25; DB 3; Length 172;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTACTTAAATATATAAATTAGCTG 2036  
DB 143 CTACTTAAATATATAAATTAGCTG 167  
  
RESULT 17  
AAL04285  
ID AAL04285 standard; DNA; 213 BP.  
XX  
AC AAL04285;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human reproductive system related antigen DNA SEQ ID NO: 6973.  
XX  
XX Human; reproductive system related antigen; reproductive system disorder;  
KM cancer; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JUN-2001; 2001WO-US001339.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214866P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220966P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 23-AUG-2000; 2000US-0227089P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0234240P.  
PR 14-SEP-2000; 2000US-0234241P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236376P.  
PR 29-SEP-2000; 2000US-0236377P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 03-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.



PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249257P.  
 PR 17-NOV-2000; 2000US-0249258P.  
 PR 17-NOV-2000; 2000US-0249259P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251899P.  
 PR 11-DEC-2000; 2000US-0251909P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-465570/50.  
 DR  
 XX  
 PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
 XX used in preventing, treating or ameliorating a medical condition.  
 PS Disclosure; SEQ ID NO 6973; 1297bp + Sequence Listing; English.  
 XX  
 CC The present invention provides the protein and coding sequences of a  
 CC number of human reproductive system related antigens. These can be used  
 CC in the prevention and treatment of reproductive system disorders,  
 CC including cancer. The present sequence is a genomic sequence encoding a  
 CC protein of the invention  
 XX  
 SQ Sequence 213 BP; 57 A; 46 C; 64 G; 46 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 4; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 0.4;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 DB 102 CTACTAAAAATATATAAATTAGCTG 126  
 RESULT 18  
 AAC20108  
 ID AAC20108 standard; cDNA; 239 BP.  
 XX  
 AC AAC20108;  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX  
 DE Human secreted protein 5' EST, SEQ ID NO: 24183.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KM gene therapy; chromosome mapping; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1033401-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 21-FEB-2000; 2000EP-00200610.  
 XX  
 PR 26-FEB-1999; 99US-0122487P.  
 XX  
 PA (GSEST ) GENSET.  
 PI Dumas Mline Edwards J, Duclert A, Giordano J;  
 XX WPI; 2000-500381/45.  
 DR  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX  
 PS Claim 1; SEQ ID NO 24183; 71pp + Sequence Listing; English.  
 XX  
 CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors  
 XX  
 SQ Sequence 239 BP; 52 A; 59 C; 66 G; 62 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 3; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 0.4;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 DB 175 CTACTAAAAATATATAAATTAGCTG 199  
 RESULT 19  
 AAAS7008  
 ID AAAS7008 standard; cDNA; 268 BP.  
 XX  
 AC AAAS7008;  
 XX

DT 14-NOV-2000 (first entry)  
XX Human colon cancer cell CDNA sequence #136.  
DE  
XX  
XX Human; arbitrary primer; CDNA synthesis; config sequence construction;  
KW open reading frame; ORF; low stringency; CDNA sequencing; ss.  
XX  
XX Homo sapiens.  
XX  
PN MO200031299-A2.  
XX  
PD 02-JUN-2000.  
XX  
XX 19-NOV-1999; 99WO-US027430.  
PF  
XX 20-NOV-1998; 98US-00196716.  
PR  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX  
XX Simpson AJG, Dias Neto E, Brentani RR;  
PI  
XX WPI; 2000-400100/34.  
DR  
XX  
XX  
PT Method for determining open reading frames of the genome of an organism  
PT using primers at low stringency conditions, useful in the construction of  
PT configs or constructs of sequenced nucleic acid molecules.  
XX  
XX Example 6; Page 52; 113pp; English.  
XX  
XX The present sequence is a cDNA sequence obtained using a method for  
CC determining open reading frames (ORFs) of the genome of an organism. An  
CC aliquot of mRNA from human colon cancer cells was mixed with a single,  
CC arbitrary primer, Moloney murine leukemia virus reverse transcriptase,  
CC reverse transcriptase buffer and dNTPs. The mixture was incubated under  
CC low stringency conditions to yield single stranded cDNA. The same primer  
CC was then used to amplify the cDNA by PCR. Rather than providing  
CC nucleotide sequence information from the non-coding terminus of nucleic  
CC acid molecules, the method provides information on the more interesting  
CC and relevant internal portions, such as ORFs. The method also permits the  
CC construction of configs of sequenced nucleic acid molecules  
XX  
SQ Sequence 268 BP; 81 A; 54 C; 78 G; 47 T; 0 U; 8 Other;  
XX  
Query Match 1.2%; Score 25; DB 3; Length 268;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 63 CTACTAAAAATATATAAATTAGCTG 87  
XX  
RESULT 20  
ABT12430  
XX ID ABT12430 standard; DNA; 268 BP.  
XX  
XX  
AC ABT12430;  
XX  
DT 02-JAN-2003 (first entry)  
XX  
XX Orestes sequencing method related DNA SEQ ID No 136.  
DE  
XX  
XX Open reading frame; ORF; low stringency; sequencing; ORESTES; bronchitis;  
KW identification; internal region; population study; heart disease;  
KW Alzheimer's disease; autoimmune disease; congenital disease; human; ds.  
XX  
XX Homo sapiens.  
XX  
XX  
XX MO200274994-A2.  
XX  
XX 26-SEP-2002.  
PD  
XX  
PF 01-NOV-2001; 2001WO-US046665.  
XX

XX  
PR 07-NOV-2000; 2000US-0246313P.  
XX  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX  
XX Simpson AJG, Dias-Neto E, Brentani RR;  
PI  
XX WPI; 2002-750567/81.  
DR  
XX  
XX  
PT Determining ORFs of the genome by contacting mRNA with a single  
PT oligonucleotide primer, useful for population studies that identify genes  
PT associated with diseases such as cardiovascular, autoimmune and  
PT congenital diseases.  
XX  
XX Example 6; Page 81; 166pp; English.  
XX  
XX The invention relates to a novel method for determining open reading  
CC frames (ORFs) of the genome of an organism comprising contacting  
CC messenger RNA from a cell of the organism with a single, oligonucleotide  
CC primer at low stringency. The method is an improved ORESTES sequencing  
CC method. The methods of the present invention are useful for sequencing  
CC of, essentially, the entire coding component of an organism,  
CC identification of previously unknown nucleic acid molecules, assigning  
CC previously identified nucleotide sequences to internal regions of genes  
CC and population studies that identify genes associated with diseases such  
CC as heart disease, bronchitis, Alzheimer's disease, autoimmune diseases  
CC and congenital diseases. This polynucleotide represents a DNA sequence  
CC that is used in the novel method of the invention  
XX  
SQ Sequence 268 BP; 81 A; 54 C; 78 G; 47 T; 0 U; 8 Other;  
XX  
Query Match 1.2%; Score 25; DB 6; Length 268;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 63 CTACTAAAAATATATAAATTAGCTG 87  
XX  
RESULT 21  
ACD91724  
XX ID ACD91724 standard; CDNA; 268 BP.  
XX  
AC ACD91724;  
XX  
DT 23-SEP-2003 (first entry)  
XX  
XX Human colon cancer cell expressed cDNA #136.  
DE  
XX  
XX Open reading frame detection; genome sequencing; colon cancer;  
KW breast cancer; population genome analysis; genetic shift; cancer;  
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;  
KW agriculture; food crop genome; resistance gene; retrovirus;  
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;  
KW gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX  
XX US2002155438-A1.  
XX  
XX 24-OCT-2002.  
PD  
XX  
XX 27-SEP-1999; 99US-00406117.  
PF  
XX 20-NOV-1998; 98US-00196716.  
PR  
XX  
XX (SIMP/) SIMPSON A J G.  
XX (NETO/) NETO E D.  
XX (BRENT/) BRENTANI R R.  
XX  
XX Simpson AJG, Neto ED, Brentani RR;  
PI  
XX

DR WPI; 2003-182626/18.  
XX  
PT Determining open reading frames of genome of an organism e.g. a human  
PT suffering from cancer involves use of single oligonucleotide primer at  
PT low stringency for preparing single-stranded cDNA from mRNA of  
PT individual.  
PS Example 6; Page 22; 959pp; English.  
XX  
XX The invention describes a method of determining open reading frames in  
CC the genome of organism, comprising contacting mRNA from cell of organism  
CC with a single oligonucleotide primer (I) at low stringency, preparing  
CC single-stranded cDNA by reverse transcribing mRNA with (II), amplifying  
CC cDNA, sequencing the product, and repeating the contacting, preparing  
CC and amplifying steps with different primers and sequencing resulting  
CC nucleic acids. The method is useful for: determining that a known  
CC nucleotide sequence from a genome of an organism corresponds to a contig,  
CC nucleotide sequence of an open reading frame; for preparing a contig,  
CC nucleic acid molecule from a genome of an organism; and for sequencing  
CC all or part of a genome of an organism. mRNA is obtained from mammalian  
CC or human cell which is associated with a pathological condition e.g. a  
CC colon cancer or breast cancer cell. The method is useful for analyses of  
CC populations of subjects and can be used to carry out genetic analyses of  
CC large or small populations. further, it can be used to study living  
CC systems to determine if, e.g. there have been genetic shifts which render  
CC an individual or population more or less likely to be afflicted with  
CC diseases such as cancer, to determine antibiotic resistance or non-  
CC tolerance, and so forth. The method can also be used in the study of  
CC congenital diseases, and the risk of affliction to a fetus, as well as  
CC the study of whether the conditions are likely to be passed to offspring  
CC through ova or sperm. The analyses for pathological conditions can be  
CC carried out in all animals, plants, birds, fish, etc. Using this method,  
CC in the area of agriculture, for example the genomes of food crops can be  
CC studied to determine if resistance genes are present, defects in plant  
CC genomes can also be studied in this way. Similarly, the method permits  
CC determination of the pathogens which integrate into the genome, such as  
CC retroviruses and other integrating viruses such as influenza virus, have  
CC undergone shifts or mutations, which may require different approaches to  
CC therapy. This method is also applied to eukaryotic pathogens, such as  
CC trypanosomes, different types of Plasmodium, etc. The method essentially  
CC eliminates sequencing of non-coding portions. This sequence represents a  
CC polynucleotide isolated from human colon cancer cell cDNA library  
XX  
SQ Sequence 268 BP; 81 A; 54 C; 78 G; 47 T; 0 U; 8 Other;  
Query Match 1.2%; Score 25; DB 10; Length 268;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTAAGTAAATATTAATTAAGCTG 2036  
DB 63 CTAAGTAAATATTAATTAAGCTG 87  
RESULT 22  
ID AAK83336/c  
ID AAK83336 standard; DNA; 274 BP.  
XX  
XX AAK83336;  
AC  
XX  
XX 07-NOV-2001 (first entry)  
DT  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38148.  
DE  
XX  
XX Human immune/haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157182-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX

PF 17-JAN-2001; 2001WO-US001354.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205155P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225265P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229503P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-023400P.  
PR 14-SEP-2000; 2000US-023401P.  
PR 14-SEP-2000; 2000US-023401P.  
PR 14-SEP-2000; 2000US-023401P.  
PR 14-SEP-2000; 2000US-023401P.  
PR 14-SEP-2000; 2000US-023401P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234984P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.

PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250191P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	11-DEC-2000;	2000US-0251990P.
PR	05-JAN-2001;	2001US-02559678P.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
XX		
XX	WPI; 2001-483426/52.	
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides	

Query Match	1.2%; Score 25; DB 4; Length 274;
Best Local Similarity	100.0%; Pred. No. 0.4;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	2012 CTACTAAATAATATTAATAATTGCTG 2036
Db	70 CTACTAAATAATATTAATAATTGCTG 46
RESULT 23	
AAK75842/c	
ID	AAK75842 standard; DNA; 282 BP.
XX	
AC	AAK75842;
XX	
DT	07-NOV-2001 (first entry)
XX	
DB	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30654.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KV	Cyclostatic; gene therapy; vaccine; metastasis; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US001354.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218290P.
PR	26-JUL-2000; 2000US-0220963P.
PR	14-AUG-2000; 2000US-0224518P.

PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246610P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246611P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	23-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	30-AUG-2000;	2000US-0237009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	01-SEP-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0239509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0239513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249246P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249267P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-0232339P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232339P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232340P.	PR	06-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251869P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234274P.	PR	08-DEC-2000;	2000US-0251990P.
PR	25-SEP-2000;	2000US-0234997P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234998P.	PR	05-JAN-2001;	2001US-0255678P.
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235836P.	PI	WPI; 2001-483426/52.	
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	DR		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	XX		
PR	29-SEP-2000;	2000US-0236370P.	PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	02-OCT-2000;	2000US-0237037P.	PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	XX	Disclosure; SEQ ID NO 30654;	

XX Sequence 282 BP; 56 A; 73 C; 59 G; 94 T; 0 U; 0 Other;  
 SQ Query Match 1.2%; Score 25; DB 4; Length 282;  
 Best Local Similarity 100.0%; Pred.No. 0.4;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||||  
 Db 263 CTACTAAAAATATATAAATTAGCTG 239

RESULT 24  
 ABL77830  
 ID ABL77830 standard; cDNA; 282 BP.  
 XX ABL77830;  
 AC 17-MAY-2002 (first entry)  
 XX  
 DT Human ovarian cancer related cDNA clone SEQ ID NO:808.  
 XX  
 DE Human; ovarian cancer; ovarian tumour; cyrostatic; gene; se.  
 XX  
 KM Homo sapiens.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200192581-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 29-MAY-2001; 2001WO-US017756.  
 XX  
 PR 26-MAY-2000; 2000US-0207484P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Algate PA, Harlocker SL, Jones R;  
 XX  
 DR WPI; 2002-122075/16.  
 XX  
 PT Composition for therapy and diagnosis of ovarian cancer comprising  
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
 PT polypeptide, antibody specific to polypeptide or T cell expressing  
 PT polypeptide.  
 XX  
 PS Claim 1; SEQ ID NO 808; 489pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising: carriers  
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI)  
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
 CC or antigen presenting cells that express (II). (I) has cyrostatic  
 CC activity. An oligonucleotide (IV) that hybridises to (SI) can be used for  
 CC detecting ovarian cancer in a patient's biological sample preferably  
 CC serum or ovarian tissue. The method comprises contacting a biological  
 CC sample from a patient with (IV), detecting the amount of polynucleotide  
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
 CC value and thereby detecting ovarian cancer in the patient, where the  
 CC amount of polynucleotide hybridising to (IV) is detected preferably by  
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
 CC useful for stimulating and/or expanding T cells specific for an ovarian  
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
 CC useful in design and preparation of ribozyme molecules for inhibiting  
 CC expression of the tumour polypeptides and proteins in tumour cells; and  
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
 CC library using well known techniques  
 XX  
 SQ Sequence 282 BP; 70 A; 58 C; 86 G; 67 T; 0 U; 1 Other;

Query Match 1.2%; Score 25; DB 6; Length 282;  
 Best Local Similarity 100.0%; Pred.No. 0.4;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||||  
 Db 99 CTACTAAAAATATATAAATTAGCTG 123

RESULT 25  
 AAV88195/c  
 ID AAV88195 standard; cDNA; 292 BP.  
 XX AAV88195;  
 AC 12-FEB-1999 (first entry)  
 XX  
 DT EST clone FX55.  
 XX  
 DE  
 XX  
 KM Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 KM tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 KM chemotaxis; chemokinesis; haemostasis; gene therapy; chromolysis;  
 KM receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9845437-A2.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US006956.  
 XX  
 PR 10-APR-1997; 97US-00837312.  
 XX  
 PA (GEMV ) GENETICS INST INC.  
 XX  
 PI Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;  
 PI Spaulding V, Agosto MJ;  
 XX  
 DR WPI; 1999-070078/06.  
 XX  
 PT New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.  
 XX  
 PS Claim 1; Page 315; 641pp; English.  
 XX  
 CC The present sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 292 BP; 57 A; 76 C; 64 G; 95 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 2; Length 292;  
 Best Local Similarity 100.0%; Pred.No. 0.4;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||||  
 Db 264 CTACTAAAAATATATAAATTAGCTG 240

RESULT 26  
 AAK76822  
 ID AAK76822 standard; DNA; 302 BP.  
 XX

AC AAK76822;  
XX 07-NOV-2001 (first entry)  
XX Human immune/haematopoietic antigen genomic sequence SRQ ID NO:31634.  
DE Human immune/haematopoietic antigen genomic sequence SRQ ID NO:31634.  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
XX MO200157182-A2.  
XX 09-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001354.  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 14-AUG-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225477P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 14-AUG-2000; 2000US-0225759P.  
XX 18-AUG-2000; 2000US-0226279P.  
XX 22-AUG-2000; 2000US-0226681P.  
XX 22-AUG-2000; 2000US-0226688P.  
XX 22-AUG-2000; 2000US-0227182P.  
XX 23-AUG-2000; 2000US-0227009P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 06-SEP-2000; 2000US-0230437P.  
XX 06-SEP-2000; 2000US-0230438P.  
XX 08-SEP-2000; 2000US-0231422P.  
XX 08-SEP-2000; 2000US-0231433P.  
XX 08-SEP-2000; 2000US-0231444P.  
XX 08-SEP-2000; 2000US-0231413P.  
XX 08-SEP-2000; 2000US-0231414P.  
XX 08-SEP-2000; 2000US-0232080P.  
XX 08-SEP-2000; 2000US-0232081P.  
XX 12-SEP-2000; 2000US-0231968P.  
XX 14-SEP-2000; 2000US-0232397P.  
XX 14-SEP-2000; 2000US-0232398P.  
XX 14-SEP-2000; 2000US-0232399P.  
XX 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 25-SEP-2000; 2000US-0234999P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 17-NOV-2000; 2000US-0249301P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251889P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-483426/52.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
 PS  
 XX Disclosure; SEQ ID NO 31634; 3071pp + Sequence listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention  
 CC  
 XX  
 SQ Sequence 302 BP; 85 A; 71 C; 93 G; 53 T; 0 U; 0 Other;  
 Query Match 1.2%; Score 25; DB 4; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 0.4;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 Db 113 CTACTAAATATATAAATTAGCTG 137  
 RESULT 27  
 AAF65644  
 ID AAF65644 standard; cDNA; 357 BP.  
 XX  
 AC AAF65644;  
 XX  
 DT 09-APR-2001 (first entry)  
 XX  
 DE Novel human polynucleotide, SEQ ID NO: 1400.  
 XX  
 KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;  
 KW breast cancer; lung cancer; cancer detection; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200102568-A2.  
 XX  
 PD 11-JAN-2001.  
 XX  
 PF 30-JUN-2000; 2000WO-US018374.  
 XX  
 PR 02-JUL-1999; 99US-0142310P.  
 PR 02-JUL-1999; 99US-0142311P.  
 XX  
 PA (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.  
 XX  
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J;  
 PI Kassam A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;  
 PI Drmanac R, Chtenjakov R, Drmanac S, Dickson M, Labat I;  
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;  
 XX  
 DR WPI; 2001-091805/10.  
 XX  
 PT Library of polynucleotides for diagnosing a cancerous state of a  
 PT mammalian cell and detecting cancer, particularly of the colon or  
 PT prostate, comprises 3351 human polynucleotide sequences.  
 XX  
 PS Claim 9; Page 745; 1046pp; English.  
 XX  
 CC The present sequence is one of 3351 sequences in a library of human  
 CC polynucleotides. The library is used to detect differentially expressed  
 CC genes correlated with a cancerous state of a mammalian cell and can  
 CC detect colon, prostate, breast and lung cancer. The library can be used  
 CC to produce probes for detection of mRNA and to produce additional copies  
 CC of the polynucleotides. The probes can be used for chromosome mapping of  
 CC the polynucleotide and for detection of transcription levels. Ribozymes  
 CC or antisense oligonucleotides can be generated. The polynucleotides and  
 CC their gene products are used as genetic or biochemical markers (e.g. in  
 CC blood or tissues) that will detect the earliest changes along the  
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and  
 CC preventive interventions. The polynucleotides, polypeptides and  
 CC antibodies against them can be used in pharmaceutical compositions to  
 CC treat the cancers and proliferative disorders such as neoplasia,  
 CC dysplasia and hyperplasia  
 CC  
 XX  
 SQ Sequence 357 BP; 100 A; 69 C; 84 G; 104 T; 0 U; 0 Other;  
 Query Match 1.2%; Score 25; DB 5; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 Db 329 CTACTAAATATATAAATTAGCTG 353  
 RESULT 28  
 AAK78393  
 ID AAK78393 standard; DNA; 375 BP.  
 XX  
 AC AAK78393;  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33205.  
 XX  
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157182-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001354.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.





CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 375 BP; 129 A; 79 C; 102 G; 65 T; 0 U; 0 Other;  
  
Query Match 1.2%; Score 25; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTAAGTAAATATATATATATAGCTG 2036  
DB 196 CTAAGTAAATATATATATATAGCTG 220  
  
RESULT 29  
AAK78396  
ID AAK78396 standard; DNA; 375 BP.  
XX  
AC AAK78396;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33208.  
XX  
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
PD  
XX 09-AUG-2001.  
XX  
PF 17-JUN-2001; 2001WO-US001354.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227709P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244674P.  
PR 08-NOV-2000; 2000US-0244752P.  
PR 08-NOV-2000; 2000US-0244766P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.



PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234977P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236377P.  
 PR 29-SEP-2000; 2000US-0236388P.  
 PR 29-SEP-2000; 2000US-0236399P.  
 PR 29-SEP-2000; 2000US-0236399P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241212P.  
 PR 20-OCT-2000; 2000US-0241215P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PI  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-483426/52.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
 XX  
 PS Disclosure; SEQ ID NO 35941; 3071bp + Sequence Listing; English.  
 XX  
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 XX Sequence 389 BP; 85 A; 87 C; 70 G; 147 T; 0 U; 0 Other;  
 SQ  
 Query Match 1.2%; Score 25; DB 4; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTACTTAAAAATATATAAATTACCTG 2036  
 DB 246 CTACTTAAAAATATATAAATTACCTG 222  
 RESULT 31  
 AAK81130/c  
 ID AAK81130 standard; DNA; 389 BP.  
 XX  
 AC AAK81130;  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35942.  
 XX  
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KM cytostatic; gene therapy; vaccine; metastasis; ds.  
 OS Homo sapiens.  
 OS  
 PN WO200157182-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001354.  
 XX

PR	31-JAN-2000	2000US-0179065P	PR	02-OCT-2000	2000US-0237037P
PR	04-FEB-2000	2000US-0180628P	PR	02-OCT-2000	2000US-0237038P
PR	24-FEB-2000	2000US-0184664P	PR	02-OCT-2000	2000US-0237039P
PR	02-MAR-2000	2000US-0186350P	PR	02-OCT-2000	2000US-0237040P
PR	16-MAR-2000	2000US-0189874P	PR	13-OCT-2000	2000US-0239335P
PR	17-MAR-2000	2000US-0190076P	PR	13-OCT-2000	2000US-0239937P
PR	18-APR-2000	2000US-0198123P	PR	20-OCT-2000	2000US-0240960P
PR	19-MAY-2000	2000US-0205515P	PR	20-OCT-2000	2000US-0241221P
PR	07-JUN-2000	2000US-0209467P	PR	20-OCT-2000	2000US-0241785P
PR	28-JUN-2000	2000US-0214868P	PR	20-OCT-2000	2000US-0241786P
PR	30-JUN-2000	2000US-0215135P	PR	20-OCT-2000	2000US-0241787P
PR	07-JUL-2000	2000US-0216647P	PR	20-OCT-2000	2000US-0241808P
PR	07-JUL-2000	2000US-0216880P	PR	20-OCT-2000	2000US-0241809P
PR	11-JUL-2000	2000US-0217487P	PR	20-OCT-2000	2000US-0241826P
PR	14-JUL-2000	2000US-0218290P	PR	01-NOV-2000	2000US-0244617P
PR	26-JUL-2000	2000US-0218296P	PR	08-NOV-2000	2000US-0246474P
PR	26-JUL-2000	2000US-0220966P	PR	08-NOV-2000	2000US-0246475P
PR	14-AUG-2000	2000US-0224518P	PR	08-NOV-2000	2000US-0246476P
PR	14-AUG-2000	2000US-0224519P	PR	08-NOV-2000	2000US-0246477P
PR	14-AUG-2000	2000US-0225213P	PR	08-NOV-2000	2000US-0246478P
PR	14-AUG-2000	2000US-0225214P	PR	08-NOV-2000	2000US-0246523P
PR	14-AUG-2000	2000US-0225266P	PR	08-NOV-2000	2000US-0246524P
PR	14-AUG-2000	2000US-0225267P	PR	08-NOV-2000	2000US-0246525P
PR	14-AUG-2000	2000US-0225268P	PR	08-NOV-2000	2000US-0246526P
PR	14-AUG-2000	2000US-0225270P	PR	08-NOV-2000	2000US-0246527P
PR	14-AUG-2000	2000US-0225447P	PR	08-NOV-2000	2000US-0246528P
PR	14-AUG-2000	2000US-0225757P	PR	08-NOV-2000	2000US-0246532P
PR	14-AUG-2000	2000US-0225758P	PR	08-NOV-2000	2000US-0246609P
PR	14-AUG-2000	2000US-0225759P	PR	08-NOV-2000	2000US-0246610P
PR	18-AUG-2000	2000US-0226279P	PR	08-NOV-2000	2000US-0246611P
PR	22-AUG-2000	2000US-0226681P	PR	17-NOV-2000	2000US-0246613P
PR	22-AUG-2000	2000US-0226868P	PR	17-NOV-2000	2000US-0249207P
PR	22-AUG-2000	2000US-0227182P	PR	17-NOV-2000	2000US-0249208P
PR	23-AUG-2000	2000US-02277009P	PR	17-NOV-2000	2000US-0249209P
PR	30-AUG-2000	2000US-0228924P	PR	17-NOV-2000	2000US-0249210P
PR	01-SEP-2000	2000US-0229287P	PR	17-NOV-2000	2000US-0249211P
PR	01-SEP-2000	2000US-0229343P	PR	17-NOV-2000	2000US-0249212P
PR	01-SEP-2000	2000US-0229344P	PR	17-NOV-2000	2000US-0249213P
PR	01-SEP-2000	2000US-0229345P	PR	17-NOV-2000	2000US-0249214P
PR	05-SEP-2000	2000US-0229509P	PR	17-NOV-2000	2000US-0249215P
PR	05-SEP-2000	2000US-0229513P	PR	17-NOV-2000	2000US-0249216P
PR	06-SEP-2000	2000US-0230437P	PR	17-NOV-2000	2000US-0249217P
PR	06-SEP-2000	2000US-0230438P	PR	17-NOV-2000	2000US-0249218P
PR	08-SEP-2000	2000US-0231242P	PR	17-NOV-2000	2000US-0249244P
PR	08-SEP-2000	2000US-0231243P	PR	17-NOV-2000	2000US-0249245P
PR	08-SEP-2000	2000US-0231244P	PR	17-NOV-2000	2000US-0249264P
PR	08-SEP-2000	2000US-0231413P	PR	17-NOV-2000	2000US-0249265P
PR	08-SEP-2000	2000US-0231414P	PR	17-NOV-2000	2000US-0249297P
PR	08-SEP-2000	2000US-0232080P	PR	17-NOV-2000	2000US-0249299P
PR	08-SEP-2000	2000US-0232081P</			

PS Disclosure; SEQ ID NO 35942; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 389 BP; 85 A; 87 C; 70 G; 147 T; 0 U; 0 Other;  
  
Query Match 1.2%; Score 25; DB 4; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTAAGTAAATATATAAATTAGCTG 2036  
DB 246 CTAAGTAAATATATAAATTAGCTG 222  
|||||  
  
RESULT 32  
AAK96262/c  
ID AAK96262 standard; DNA; 401 BP.  
XX  
AC AAK96262;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Human neuregulin gene single nucleotide polymorphism SNP8NRG1100717.  
XX  
KW Human; neuregulin-1 associated gene 1; NRG1AG1; Schizophrenia gene;  
KW gene therapy; single nucleotide polymorphism; SNP; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200164876-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US006376.  
XX  
PR 28-FEB-2000; 2000US-00515715.  
XX  
PA (DECO-) DECODE GENETICS EHF.  
XX  
PI Stefansson H, Steinthorsdottir V, Gulcher JR;  
XX WPI; 2001-550179/61.  
XX  
PT Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for  
XX preventing diagnosing and treating schizophrenia.  
XX  
PS Disclosure; Page 681; 750bp; English.  
XX  
XX This sequence represents a single nucleotide polymorphism (SNP) of the  
CC human neuregulin-1 associated gene 1 (NRG1AG1) of the invention. The  
CC NRG1AG1 gene is also referred to as the human Schizophrenia gene. The  
CC invention also relates to fragments or variants of the gene and the  
CC NRG1AG1 polypeptides they encode. The NRG1AG1 nucleic acids and the  
CC polypeptides may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate NRG1AG1 expression. For example,  
CC they may be used to treat disorders associated with decreased expression

CC by rectifying mutations or deletions in a patient's genome that affect  
CC the activity of NRG1AG1 by expressing inactive proteins or to supplement  
CC the patient's own production of NRG1AG1. Additionally, the gene may be  
CC used to produce NRG1AG1 polypeptides, by inserting the nucleic acids into  
CC a host cell and culturing the cell to express the protein. The gene may  
CC also be used as DNA probes and primers in diagnostic assays to detect and  
CC quantitate the presence of similar nucleic acids in samples, and  
CC therefore which patients may be in need of restorative therapy. The  
CC NRG1AG1 polypeptides may also be used as antigens in the production of  
CC antibodies against NRG1AG1 and in assays to identify modulators of  
CC NRG1AG1 expression and activity. Anti-NRG1AG1 antibodies and antagonists  
CC may also be used to down regulate expression and activity. Anti-NRG1AG1  
CC antibodies may also be used as diagnostic agents for detecting the  
CC presence of NRG1AG1 polypeptides in samples. NRG1AG1 is associated with  
CC schizophrenia which may be prevented, diagnosed and/or treated by the  
CC above methods  
XX  
SQ Sequence 401 BP; 78 A; 93 C; 99 G; 130 T; 0 U; 1 Other;  
  
Query Match 1.2%; Score 25; DB 4; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTAAGTAAATATATAAATTAGCTG 2036  
DB 184 CTAAGTAAATATATAAATTAGCTG 160  
|||||  
  
RESULT 33  
AAK97755/c  
ID AAK97755 standard; DNA; 401 BP.  
XX  
AC AAK97755;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Human neuregulin gene single nucleotide polymorphism SNP8NRG1100717.  
XX  
KW Human; neuregulin 1 gene; schizophrenia; gene therapy; SNP;  
KW single nucleotide polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200164877-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US006377.  
XX  
PR 28-FEB-2000; 2000US-00515716.  
XX  
PA (DECO-) DECODE GENETICS EHF.  
XX  
PI Stefansson H, Steinthorsdottir V, Gulcher JR;  
XX WPI; 2001-514841/56.  
XX  
PT Neuregulin 1 nucleic acids and proteins useful for diagnosing preventing  
XX and treating schizophrenia.  
XX  
PS Disclosure; Page 266; 756bp; English.  
XX  
XX This sequence represents a single nucleotide polymorphism (SNP) from the  
CC human neuregulin 1 gene of the invention. The invention also relates to  
CC fragments or variants of the neuregulin 1 gene. The gene and its proteins  
CC may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate neuregulin 1 expression, such as  
CC schizophrenia. For example they may be used to treat disorders associated  
CC with decreased neuregulin 1 expression by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of neuregulin 1  
CC by expressing inactive proteins or to supplement the patient's own  
CC production of polypeptides. Additionally, the gene may be used to produce  
CC the neuregulin 1 protein, by inserting the nucleic acids into a host cell

CC and culturing the cell to express the protein. The gene and its  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. The protein may also be used as antigens in the production of  
 CC antibodies against neuregulin 1 and in assays to identify modulators of  
 CC neuregulin 1 expression and activity. The antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The antibodies may  
 CC also be used as diagnostic agents for detecting the presence of  
 CC neuregulin 1 in samples

SQ Sequence 401 BP; 78 A; 93 C; 99 G; 130 T; 0 U; 1 Other;

Query Match 1.2%; Score 25; DB 4; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAAAATATTAATAATTAGCTG 2036  
 |||||  
 DB 184 CTACTAAAAATATTAATAATTAGCTG 160

RESULT 34  
 ABT01032/c  
 ID ABT01032 standard; DNA; 401 BP.

AC ABT01032;  
 DT 07-NOV-2002 (first entry)

DE Human neuregulin-1-associated gene 1 fragment SEQ ID NO: 1061.

KW Human; neuregulin 1; neuregulin-1-associated gene 1; NRGI; NRGIAG1;  
 KW schizophrenia; chromosome 8p12; single nucleotide polymorphism; SNP;  
 KW neuroleptic; gene therapy; splice variant; gene; ds.

OS Homo sapiens.

PN US2002045577-A1.

PD 18-APR-2002.

PF 28-FEB-2001; 2001US-00795686.

PR 28-FEB-2000; 2000US-00515716.

PA (DECO-) DECODE GENETICS EHF.

PI Stefansson H, Steinhorstodttr V, Gulcher JR;

DR WPI; 2002-425447/45.

PT New neuregulin 1 gene, schizophrenia gene residing in 1.5 Mb segment on  
 PT human chromosome 8p12, useful for diagnosing susceptibility to or  
 PT treating schizophrenia and for screening schizophrenia treating agents.

PS Disclosure; Page 622; 700bp; English.

XX The present invention relates to the human neuregulin 1 gene (NRGI),  
 CC single nucleotide polymorphisms within which were identified as being  
 CC associated with an increased susceptibility to schizophrenia, which is  
 CC located on chromosome 8p12. Also found within the same sequence is the  
 CC neuregulin-1-associated gene 1 (NRGIAG1). The gene is useful for treating  
 CC schizophrenia in an individual, for diagnosing susceptibility to  
 CC schizophrenia, and for screening for agents useful in the treatment of  
 CC the disease. The present sequence is a fragment of the NRGIAG1 gene of  
 CC the invention containing a polymorphic site

SQ Sequence 401 BP; 78 A; 93 C; 99 G; 130 T; 0 U; 1 Other;

Query Match 1.2%; Score 25; DB 6; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAAAATATTAATAATTAGCTG 2036  
 |||||  
 DB 184 CTACTAAAAATATTAATAATTAGCTG 160

RESULT 35  
 ABT02525/c  
 ID ABT02525 standard; DNA; 401 BP.

AC ABT02525;

DT 07-NOV-2002 (first entry)

DE Human neuregulin-1-associated gene 1 fragment SEQ ID NO: 1061.

KW Human; neuregulin 1; neuregulin-1-associated gene 1; NRGI; NRGIAG1;  
 KW schizophrenia; chromosome 8p12; single nucleotide polymorphism; SNP;  
 KW neuroleptic; gene therapy; splice variant; gene; ds.

OS Homo sapiens.

PN US2002094954-A1.

PD 18-JUL-2002.

PF 28-FEB-2001; 2001US-00795686.

PR 28-FEB-2000; 2000US-00515715.

PA (DECO-) DECODE GENETICS EHF.

PI Stefansson H, Steinhorstodttr V, Gulcher JR;

DR WPI; 2002-665799/71.

PT NRGIAG1, useful for treating or diagnosing susceptibility to  
 PT schizophrenia, or for assaying a sample for the presence of NRGIAG1  
 PT nucleic acid.

PS Disclosure; Page 621; 700bp; English.

XX The present invention relates to the human neuregulin-1-associated gene  
 CC 1 (NRGIAG1), single nucleotide polymorphisms within which were identified  
 CC as being associated with an increased susceptibility to schizophrenia,  
 CC which is located on chromosome 8p12. Also found within the same sequence  
 CC is the neuregulin 1 gene (NRGI). The gene is useful for treating  
 CC schizophrenia in an individual, for diagnosing susceptibility to  
 CC schizophrenia, and for screening for agents useful in the treatment of  
 CC the disease. The present sequence is a fragment of the NRGIAG1 gene of  
 CC the invention containing a polymorphic site

SQ Sequence 401 BP; 78 A; 93 C; 99 G; 130 T; 0 U; 1 Other;

Query Match 1.2%; Score 25; DB 6; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAAAATATTAATAATTAGCTG 2036  
 |||||  
 DB 184 CTACTAAAAATATTAATAATTAGCTG 160

RESULT 36  
 ABV57504  
 ID ABV57504 standard; cDNA; 403 BP.

AC ABV57504;

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 57495.

XX

KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
OS Homo sapiens.  
XX  
XX WO200160860-A2.  
XX  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
XX 16-MAR-2000; 2000US-0189862P.  
XX 25-MAY-2000; 2000US-0207454P.  
XX 09-JUN-2000; 2000US-0211314P.  
XX 18-JUL-2000; 2000US-0219007P.  
XX 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 11059; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (I) is useful for: (a) assessing whether  
XX a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing  
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;  
XX (e) selecting a composition for inhibiting prostate cancer in a patient;  
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
XX determining whether prostate cancer has metastasized in a patient; (h)  
XX assessing the aggressiveness or indolence of prostate cancer in a patient  
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 403 BP; 108 A; 93 C; 126 G; 76 T; 0 U; 0 Other;  
  
Query Match 1.2%; Score 25; DB 5; Length 403;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
Db 250 CTACTAAAAATATATAAATTAGCTG 274  
  
RESULT 37  
AAK05173  
ID AAK05173 standard; DNA; 411 BP.  
XX  
XX AAK05173;  
XX  
XX 06-APR-1999 (first entry)  
XX  
XX Human MSH5 (hMSH5) gene intron 8 partial sequence.  
XX  
XX MSH5; hMSH5; DNA mismatch repair gene; chromosome segregation; meiosis;  
XX malignant; infertility; Down's syndrome; tumour; cancer; gene therapy;  
XX human; intron; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO901550-A1.  
XX  
XX 14-JAN-1999.  
XX  
XX

XX  
XX 02-JUL-1998; 98WO-US013850.  
XX  
XX PF 03-JUL-1997; 97US-0051686P.  
XX  
XX (DAND ) DNA FARMER CANCER INST INC.  
XX  
XX PI Kolodner R, Winand N;  
XX WPI; 1999-106052/09.  
XX  
XX New isolated human DNA mismatch repair gene, MSH5 - used for developing  
XX products for the diagnosis and therapy of disorders such as cancer,  
XX infertility and Down's syndrome.  
XX  
XX Disclosure; Page 63; 114pp; English.  
XX  
XX The invention relates to a human DNA mismatch repair gene, MSH5. Host  
XX cells containing a vector comprising the MSH5 gene is used for the  
XX recombinant production of the MSH5 protein. The MSH5 gene product is  
XX required for meiotic crossing over and segregation of chromosomes during  
XX meiosis. The products can be used for detecting an alteration in a  
XX mammalian gene as indicative of a predisposition to malignant growth of  
XX cells or indicative of a predisposition to a malady associated with  
XX inappropriate meiotic segregation such as infertility or Down's syndrome.  
XX The alterations can also be used for diagnosing a DNA mismatch pair  
XX defective tumour and for prognosis of an individual having cancer.  
XX Moreover, defects in this gene confer resistance to alkylating agents.  
XX The products can also be used to identify therapeutic agents effective  
XX against MSH5 defects and agents that affect the gene. The products can  
XX also be used for gene therapy. Sequences AAK05164-194 represent introns  
XX of the hMSH5 gene  
  
SQ Sequence 411 BP; 118 A; 76 C; 119 G; 98 T; 0 U; 0 Other;  
  
Query Match 1.2%; Score 25; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
Db 362 CTACTAAAAATATATAAATTAGCTG 386  
  
RESULT 38  
ABV17272/C  
ID ABV17272 standard; cDNA; 427 BP.  
XX  
XX ABV17272;  
XX  
XX 13-SEP-2002 (first entry)  
XX  
XX Human prostate expression marker cDNA 17263.  
XX  
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200160860-A2.  
XX  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
XX 16-MAR-2000; 2000US-0189862P.  
XX 25-MAY-2000; 2000US-0207454P.  
XX 09-JUN-2000; 2000US-0211314P.  
XX 18-JUL-2000; 2000US-0219007P.  
XX 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX



XX Schlegel R, Endege WO, Monahan JE;  
XX MPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer; stage of prostate cancer.  
XX Claim 1; Page 2863; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (1) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (1) is useful for: (a) assessing whether  
XX a patient is afflicted with prostate cancer; (b) monitoring the  
XX progression of prostate cancer in a patient; (c) assessing the efficacy  
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing  
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;  
XX (e) selecting a composition for inhibiting prostate cancer in a patient;  
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
XX determining whether prostate cancer has metastasized in a patient; (h)  
XX assessing the aggressiveness or indolence of prostate cancer in a patient  
XX ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX Sequence 427 BP; 106 A; 96 C; 86 G; 139 T; 0 U; 0 Other;  
XX  
XX Query Match 1.2%; Score 25; DB 5; Length 427;  
XX Best Local Similarity 100.0%; Pred. No. 0.39;  
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 2012 CTACTAAAAATATATAAATTAGCTG 2036  
XX DB 375 CTACTAAAAATATATAAATTAGCTG 351  
XX  
XX RESULT 39  
XX ACH47498  
XX ID ACH47498 standard; cDNA; 434 BP.  
XX  
XX ACH47498;  
XX  
XX 13-OCT-2003 (first entry)  
XX  
XX Human infant brain cDNA #1561.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRMA/) DRMANAC R T.  
XX (LABA/) LABAT I.  
XX (STAC/) STACHE-CRAIN B.  
XX (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX MPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
XX as hybridization probes, as oligomers for PCR, for chromosome and gene  
XX mapping, in the recombinant production of protein, or in generating  
XX antisense DNA or RNA.

PS Claim 1; SEQ ID NO 34710; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
XX determined by the technique of SBH (sequencing by hybridisation). Also  
XX included is a purified polypeptide comprising a sequence corresponding to  
XX a reading frame of the novel polynucleotide. The nucleic acid sequences  
XX are useful in diagnostics as expressed sequence tags (EST) for  
XX identifying expressed genes or for physical mapping of the human genome,  
XX in forensics, in assessing biodiversity, or in identifying mutations  
XX responsible for genetic disorders and other traits. The nucleotide  
XX sequences are also useful as hybridisation probes, as oligomers for PCR,  
XX for chromosome and gene mapping, in the recombinant production of  
XX protein, or in generating antisense DNA or RNA. The purified polypeptide  
XX is useful for generating antibodies specific for it. The present sequence  
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
XX for this patent did not form part of the printed specification, but was  
XX obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX Sequence 434 BP; 117 A; 110 C; 110 G; 96 T; 0 U; 1 Other;  
XX  
XX Query Match 1.2%; Score 25; DB 9; Length 434;  
XX Best Local Similarity 100.0%; Pred. No. 0.39;  
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 2012 CTACTAAAAATATATAAATTAGCTG 2036  
XX DB 171 CTACTAAAAATATATAAATTAGCTG 195  
XX  
XX RESULT 40  
XX AAH13419/C  
XX ID AAH13419 standard; cDNA; 447 BP.  
XX  
XX AAH13419;  
XX  
XX 26-JUN-2001 (first entry)  
XX  
XX Human cDNA clone (3'-primer) SEQ ID NO:10254.  
XX  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
XX Homo sapiens.  
XX  
XX EP1074617-A2.  
XX  
XX 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-00116126.  
XX  
XX 29-JUL-1999; 99JP-00248036.  
XX  
XX 27-AUG-1999; 99JP-00300253.  
XX  
XX 11-JAN-2000; 2000JP-00118776.  
XX  
XX 02-MAY-2000; 2000JP-00183767.  
XX  
XX 09-JUN-2000; 2000JP-00241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX MPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
XX length cDNAs defined in the specification, and for the detection and/or  
XX diagnosis of the abnormality of the proteins encoded by the full-length  
XX cDNAs.  
XX  
XX Claim 3; SEQ ID NO 10254; 2537bp + Sequence Listing; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602 full-  
XX length cDNAs defined in the specification. Where a primer set comprises:



PR	30-AUG-2000	2000US-0227009P
PR	01-SEP-2000	2000US-0228954P
PR	01-SEP-2000	2000US-0229287P
PR	01-SEP-2000	2000US-0229343P
PR	01-SEP-2000	2000US-0229344P
PR	01-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229559P
PR	05-SEP-2000	2000US-0229559P
PR	06-SEP-2000	2000US-0230437P
PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231143P
PR	08-SEP-2000	2000US-0231144P
PR	08-SEP-2000	2000US-0231144P
PR	08-SEP-2000	2000US-0231443P
PR	08-SEP-2000	2000US-0231443P
PR	12-SEP-2000	2000US-0231568P
PR	14-SEP-2000	2000US-0232397P
PR	14-SEP-2000	2000US-0232398P
PR	14-SEP-2000	2000US-0232399P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0233061P
PR	14-SEP-2000	2000US-0233064P
PR	14-SEP-2000	2000US-0233065P
PR	21-SEP-2000	2000US-0234223P
PR	21-SEP-2000	2000US-0234474P
PR	25-SEP-2000	2000US-0234997P
PR	25-SEP-2000	2000US-0234998P
PR	26-SEP-2000	2000US-0235844P
PR	27-SEP-2000	2000US-0235834P
PR	27-SEP-2000	2000US-0235836P
PR	29-SEP-2000	2000US-0236327P
PR	29-SEP-2000	2000US-0236567P
PR	29-SEP-2000	2000US-0236568P
PR	29-SEP-2000	2000US-0236569P
PR	29-SEP-2000	2000US-0236802P
PR	02-OCT-2000	2000US-0237037P
PR	02-OCT-2000	2000US-0237038P
PR	02-OCT-2000	2000US-0237039P
PR	02-OCT-2000	2000US-0237040P
PR	13-OCT-2000	2000US-0239335P
PR	13-OCT-2000	2000US-0239937P
PR	20-OCT-2000	2000US-0240960P
PR	20-OCT-2000	2000US-0241221P
PR	20-OCT-2000	2000US-0241785P
PR	20-OCT-2000	2000US-0241876P
PR	20-OCT-2000	2000US-0241877P
PR	20-OCT-2000	2000US-0241880P
PR	20-OCT-2000	2000US-0241889P
PR	01-NOV-2000	2000US-0244617P
PR	08-NOV-2000	2000US-0246474P
PR	08-NOV-2000	2000US-0246475P
PR	08-NOV-2000	2000US-0246476P
PR	08-NOV-2000	2000US-0246477P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246523P
PR	08-NOV-2000	2000US-0246524P
PR	08-NOV-2000	2000US-0246525P
PR	08-NOV-2000	2000US-0246526P
PR	08-NOV-2000	2000US-0246527P
PR	08-NOV-2000	2000US-0246528P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246609P
PR	08-NOV-2000	2000US-0246610P
PR	08-NOV-2000	2000US-0246611P
PR	17-NOV-2000	2000US-0249207P
PR	17-NOV-2000	2000US-0249208P
PR	17-NOV-2000	2000US-0249209P

PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0255678P.

XX	(HUMA-)	HUMAN GENOME SCI INC.
XX	PA	
XX	PI	Rosen CA, Barash SC, Ruben SM;
XX	DR	WPI, 2001-483426/52.
XX	DR	P-PsDB; AAM89773.
XX	PT	
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and metastases.	
XX		
XX	Claim 1; SEQ ID NO 7614; 3071bp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytotostic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patient's own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703	
CC	to AAK67694 represent human immune/hematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
XX	represent sequences used in the exemplification of the present invention	
XX		
XX	Sequence 453 BP; 111 A; 110 C; 117 G; 106 T; 0 U; 9 Other;	
XX		
Query Match	1.2%;	Score 25; DB 4; Length 453;
Best Local Similarity	100.0%;	Pred. No. 0.39;
Matches 25; Conservative 0;	Mismatches 0;	Indels 0; Gaps 0

QY	2012	CTACTAAAAATATAAAAATTAGCTG	2036
Db	139	CTACTAAAAATATAAAAATTAGCTG	115

RESULT 43  
ABV47067/C

```

ID  ABV47067 standard; cDNA; 467 BP.
XX
XX  ABV47067;
XX
XX  16-SEP-2002 (first entry)
XX
XX  Human prostate expression marker cDNA 47058.
XX
XX  Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX  pharmacogenomic marker; gene; ss.
XX
XX  Homo sapiens.
XX
XX  MO200160860-A2.
XX
XX  23-AUG-2001.
XX
XX  20-FEB-2001; 2001WO-US005171.
XX
XX  17-FEB-2000; 2000US-0183119P.
XX  16-MAR-2000; 2000US-0189862P.
XX  25-MAY-2000; 2000US-0207454P.
XX  09-JUN-2000; 2000US-0211314P.
XX  18-JUL-2000; 2000US-0219007P.
XX  13-DEC-2000; 2000US-0255281P.
XX
XX  (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX  Schlegel R, Endege WO, Monahan JE;
XX
XX  MPI; 2001-662795/76.
XX
XX  Novel isolated nucleic acid molecule associated with cancerous state of
XX  prostate cells and correlating with presence of prostate cancer, useful
XX  for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX  Claim 1; Page 9267; 11750pp; English.
XX
XX  The invention relates to an isolated nucleic acid molecule (I) comprising
XX  a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX  specification or its complement. (I) is useful for: (a) assessing whether
XX  a patient is afflicted with prostate cancer; (b) monitoring the
XX  progression of prostate cancer in a patient; (c) assessing the efficacy
XX  of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX  the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX  (e) selecting a composition for inhibiting prostate cancer in a patient;
XX  (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX  determining whether prostate cancer has metastasized in a patient; (h)
XX  assessing the aggressiveness or indolence of prostate cancer in a patient
XX  ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX  Sequence 467 BP; 112 A; 114 C; 97 G; 144 T; 0 U; 0 Other;
XX
XX  Query Match 1.2%; Score 25; DB 5; Length 467;
XX  Best Local Similarity 100.0%; Pred. No. 0.39;
XX  Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 2012 CTACTAAAAATATAAAAATTAGCTG 2036
XX  |||||||
XX  DB 402 CTACTAAAAATATAAAAATTAGCTG 378
XX
XX  RESULT 44
XX  ABL83732/c
XX  ID ABL83732 standard; cDNA; 472 BP.
XX
XX  ABL83732;
XX
XX  17-MAY-2002 (first entry)
XX
XX  Human ovarian cancer related cDNA clone SEQ ID NO:6710.
XX
XX  Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX

```

```

XX  Homo sapiens.
XX
XX  WO200192581-A2.
XX
XX  06-DEC-2001.
XX
XX  29-MAY-2001; 2001WO-US017756.
XX
XX  26-MAY-2000; 2000US-0207484P.
XX
XX  (CORI-) CORIXA CORP.
XX
XX  Algate PA, Harlocker SL, Jones R;
XX
XX  MPI; 2002-122075/16.
XX
XX  Composition for therapy and diagnosis of ovarian cancer comprising
XX  polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX  polypeptide, antibody specific to polypeptide or T cell expressing
XX  polypeptide.
XX
XX  Claim 1; SEQ ID NO 6710; 489pp; English.
XX
XX  The present invention describes a composition (I) comprising: carriers
XX  and immunostimulants; and a polypeptide (II) of a ovarian tumour
XX  polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
XX  from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
XX  (III) encoding (II) having a sequence (S2), a T cell population of (II),
XX  or antigen presenting cells that express (II). (I) has cytostatic
XX  activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
XX  detecting ovarian cancer in a patient's biological sample preferably
XX  serum or ovarian tissue. The method comprises contacting a biological
XX  sample from a patient with (IV), detecting the amount of polynucleotide
XX  hybridising to (IV) and comparing the amount to a predetermined cutoff
XX  value and thereby detecting ovarian cancer in the patient, where the
XX  amount of polynucleotide hybridising to (IV) is detected preferably by
XX  CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
XX  useful for stimulating and/or expanding T cells specific for an ovarian
XX  tumour protein comprising contacting T cells with (III) or (II). (III) is
XX  useful in design and preparation of ribozyme molecules for inhibiting
XX  CC expression of the tumour polypeptides and proteins in tumour cells; and
XX  CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
XX  library using well known techniques
XX
XX  Sequence 472 BP; 113 A; 133 C; 99 G; 127 T; 0 U; 0 Other;
XX
XX  Query Match 1.2%; Score 25; DB 6; Length 472;
XX  Best Local Similarity 100.0%; Pred. No. 0.39;
XX  Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 2012 CTACTAAAAATATAAAAATTAGCTG 2036
XX  |||||||
XX  DB 171 CTACTAAAAATATAAAAATTAGCTG 147
XX
XX  RESULT 45
XX  ACH15457/c
XX  ID ACH15457 standard; cDNA; 474 BP.
XX
XX  ACH15457;
XX
XX  13-OCT-2003 (first entry)
XX
XX  Human adult brain cDNA #2669.
XX
XX  Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX  genome mapping; biodiversity; genetic disorder.
XX
XX  Homo sapiens.
XX
XX  US2003073623-A1.
XX

```

PD	17-APR-2003.
PF	30-JUL-2001; 2001US-00918995.
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I.
PA	(STRAC/) STRACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
XX	
PI	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
DR	WPI; 2003-615964/58.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
XX	
PS	Claim 1; SEQ ID NO 2669; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC	determined by the technique of SBH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversity, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPRO at
CC	seqdata.uspro.gov/sequence.html?docID=20030073623
XX	
SO	Sequence 474 BP; 92 A; 131 C; 121 G; 126 T; 0 U; 4 Other;
	Query Match 1.2%; Score 25; DB 9; Length 474;
	Best Local Similarity 100.0%; Pred. No. 0.39;
	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	2012 CTACTAAATAATTAAATAATTGCTG 2036
DB	104 CTACTAAATAATTAAATAATTGCTG 80
RESULT 46	
ACH28081	
ACH28081 standard: cDNA; 486 BP.	
XX	
AC	ACH28081;
XX	
DT	13-OCT-2003 (first entry)
DE	
DE	Human adult ovary cDNA #6461.
XX	
KW	Human; BG; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX	genome mapping; biodiversity; genetic disorder.
OS	
XX	Homo sapiens.
XX	
PN	US2003073623-A1.
XX	
PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.

[illegible]

PA	(STRAC/-CREAIN B.
PA	(DICK// DICKSON M C.
PA	(JONE/) JONES L W.
XX	
PI	Dremanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
DR	WPI, 2003-615964/58.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
XX	
PS	Claim 1, SEQ ID NO 34441; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACHS0831, whose sequence was
CC	determined by the technique of SBH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversity, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?docID=20030073623
XX	
SO	Sequence 488 BP, 121 A, 122 C, 138 G, 99 T, 0 U; 8 Other;

	Query Match	Similarity	1.2%	Score 25;	DB 9;	Length 488;
	Best Local	Similarity	100.0%;	Pred. No.	0.33;	
	Matches	Conservative	0;	Mismatches	0;	Gaps 0;
OY	2012	CTACTAAATAATATAAAAAATTAGCTG	2036			
Db	232	CTACTAAAAAATATAAAAAATTAGCTG	256			

RESULT 48	
AAH13219/c	
ID	AAH13219 standard; cDNA; 496 BP.
XX	
AC	AAH13219;
DT	26-JUN-2001 (first entry)
XX	
DE	Human CDNA clone (3'-primer) SEQ ID NO:10054.
XX	
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
FN	EP1074617-A2.
XX	
PD	07-FEB-2001.
XX	
PF	28-JUL-2000; 2000EP-00116126.
XX	
PR	29-JUL-1999; 99JP-00248036.
PR	27-AUG-1999; 99JP-00300253.
PR	11-JAN-2000; 2000JP-00118776.
PR	02-MAY-2000; 2000JP-00183767.
PR	09-JUN-2000; 2000JP-00241899.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX  
DR WPI: 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.  
XX  
XX  
PS Claim 3; SEQ ID NO 10054; 2537bp + Sequence Listing; English.

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
CC represent human amino acid sequences; and AAH15629 to AAH13632 represent  
CC polynucleotides, all of which are used in the exemplification of the  
CC present invention

```
SQ Sequence 496 BP; 97 A; 127 C; 120 G; 145 T; 0 U; 7 Other;
Query Match Similarity 1.2%; Score 25; DB 4; Length 496;
      Best Local Similarity 100.08; Pied. No. 0.39;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 49	
ID	ADP90604/c
AD	ADP90604 standard; DNA; 503 BP.
XX	
AC	ADP90604;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	Human hepatic-fibrosis disease marker SEQ ID 66.
XX	
KM	Hepatic fibrosis; marker; chronic hepatitis; liver cirrhosis;
KM	hepatic carcinoma; human; ds.
XX	
OS	Homo sapiens.
XX	
PN	JP2003259877-A.
XX	
PD	16-SEP-2003.
XX	
PE	11-MAR-2002; 2002JP-00065013.
XX	
PR	11-MAR-2002; 2002JP-00065013.
XX	
PA	(SUMU ) SUMITOMO SEIYAKU KK.
XX	
WI	WI; 2003-821598/77.
XX	
PT	Hepatic fibrosis disease markers comprising polynucleotides or

PT antibodies, useful for improved diagnosis, screening and developing drugs  
PT to treat hepatitis, to control cirrhosis and carcinoma.  
PS Claim 1; SEQ ID NO 66; 313pp; Japanese.  
XX  
CC The present invention relates to hepatic-fibrosis disease markers  
CC (ADP90539-ADP90871) and related proteins (ADP90872-ADP90917). The  
CC sequences are useful for detecting and treating hepatic fibrosis caused  
CC by alcohol consumption, virus infection, etc., and the associated chronic  
CC hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The  
CC markers allow the cause of hepatic fibrosis to be clarified (diagnostic  
CC precision), so more suitable treatments can be developed and given.  
XX  
SQ Sequence 503 BP; 121 A; 117 C; 110 G; 155 T; 0 U; 0 Other;  
  
Query Match 1.2%; Score 25; DB 10; Length 503;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTAATAAAATATATAAATTAGCTG 2036  
DB 183 CTAATAAAATATATAAATTAGCTG 159  
|||||  
  
RESULT 50  
AAL15788/C  
ID AAL15788 standard; cDNA, 520 BP.  
XX  
AC AAL15788;  
XX  
DT 07-DEC-2001 (first entry)  
XX  
DE Human breast cancer expressed polynucleotide 8245.  
XX  
KW Human; breast cancer; cell marker; cytostatic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200151628-A2.  
XX  
PD 19-JUL-2001.  
XX  
PF 10-JAN-2001; 2001WO-US000798.  
XX  
PR 14-JAN-2000; 2000US-0176077P.  
XX  
PR 14-MAR-2000; 2000US-0189167P.  
XX  
PR 24-MAR-2000; 2000US-0192099P.  
XX  
PR 29-MAR-2000; 2000US-0193480P.  
XX  
PR 15-MAY-2000; 2000US-0205230P.  
XX  
PR 09-JUN-2000; 2000US-0211315P.  
XX  
PR 25-JUL-2000; 2000US-0220534P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Little J, Xu Y, Wang Y, Steinmann K;  
XX  
DR WPI; 2001-451856/48.  
XX  
PT New peptide useful as a marker for the diagnosis of breast cancer.  
XX  
PS Claim 1; Page 1490-1491; 3695pp; English.  
XX  
CC The invention relates to human breast cancer expressed polynucleotides  
CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
CC afflicted with breast cancer by examining the correlation between the  
CC expression of certain markers and the cancerous state of breast cells.  
CC The polynucleotides and encoded polypeptides are potential markers for  
CC detecting, diagnosing, monitoring, characterising treating and  
CC potentially preventing breast cancer. The polynucleotides and encoded  
CC polypeptides are also useful for isolating compounds with cytostatic  
CC activity  
XX  
SQ Sequence 520 BP; 84 A; 143 C; 130 G; 148 T; 0 U; 15 Other;

Query Match 1.2%; Score 25; DB 4; Length 520;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2012 CTAATAAAATATATAAATTAGCTG 2036  
DB 186 CTAATAAAATATATAAATTAGCTG 162  
|||||  
  
RESULT 51  
AAF94242/C  
ID AAF94242 standard; DNA; 541 BP.  
XX  
AC AAF94242;  
XX  
DT 23-MAY-2001 (first entry)  
XX  
DE Primer specific for DNA encoding secretory/membrane protein SEQ ID 676.  
XX  
KW Human; secretory protein; membrane protein; vaccine; gene therapy;  
KW rheumatoid arthritis; diabetes; PCR primer; ss.  
XX  
OS Synthetic.  
XX  
PN EP1067182-A2.  
XX  
PD 10-JAN-2001.  
XX  
PF 07-JUL-2000; 2000EP-00114090.  
XX  
PR 08-JUL-1999; 99JP-00194179.  
XX  
PR 11-JAN-2000; 2000JP-00118775.  
XX  
PR 02-MAY-2000; 2000JP-00183766.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;  
XX  
DR WPI; 2001-093989/11.  
XX  
PT Nucleic acids encoding secretory proteins/membrane proteins, useful in  
XX gene therapy or as candidate target molecules in drug development.  
XX  
PS Claim 5; SEQ ID NO 676; 609pp + Sequence Listing; English.  
XX  
CC This invention relates to nucleic acid sequences AAF931744 - AAF93916  
CC which encode human secretory or membrane proteins represented by AAB88317  
CC - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and  
CC AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the  
CC invention. The invention also includes methods for the production of  
CC antibodies directed against the proteins, and cDNA sequences, which can  
CC be used in vaccines. The polynucleotide sequences can be used in gene  
CC therapy. The polynucleotide sequences and the proteins they encode may be  
CC used in the prevention, treatment and diagnosis of diseases associated  
CC with inappropriate secretory protein/membrane protein expression. The  
CC nucleic acids and complementary sequences may also be used as DNA probes  
CC in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect  
CC and quantitate the presence of similar nucleic acid sequences in samples.  
CC They may also be used to study the expression and function of secretory  
CC proteins/membrane polypeptides and their role in metabolism. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC against them and in assays to identify modulators (agonists and  
CC antagonists) of expression and activity. The antibodies and antagonists  
CC may also be used as therapeutic agents to down regulate expression and  
CC activity. The antibodies may also be used as diagnostic agents for  
CC detecting the presence of the polypeptides in samples (e.g. by enzyme  
CC linked immunosorbent assay (ELISA). Examples of diseases which may be  
CC treated include rheumatoid arthritis and diabetes  
XX  
SQ Sequence 541 BP; 138 A; 121 C; 113 G; 160 T; 0 U; 9 Other;  
  
Query Match 1.2%; Score 25; DB 5; Length 541;

Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||  
Db 166 CTTACTAAAAATATATAAAATTAGCTG 142

## RESULT 52

ABN60792  
ID ABN60792 standard; cDNA; 550 BP.

XX  
AC ABN60792;

XX  
DT 28-JUN-2002 (first entry)

XX  
DE Human cancer related polynucleotide SEQ ID NO 759.

XX  
KW Human; cytostatic; gene expression; gene mapping; tissue profiling;

XX  
KW gene therapy; cancer; tumour; gene; ss.

XX  
OS Homo sapiens.

XX  
PN WO200214500-A2.

XX  
PD 21-FEB-2002.

XX  
PF 16-AUG-2001; 2001WO-US025840.

XX  
PR 16-AUG-2000; 2000US-0226326P.

XX  
PA (CHIR ) CHIRON CORP.

XX  
PA (HYSE-) HYSEQ INC.

XX  
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;

XX  
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;

XX  
DR MPI; 2002-241905/29.

XX  
PT New nucleic acid for producing a polypeptide, detecting differentially  
PT expressed genes correlated with a cancerous state of a mammalian cell,  
PT and inhibiting tumor growth.

XX  
XX Claim 1; SEQ ID NO 759; 883bp + Sequence Listing; English.

XX  
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)

XX  
CC with cytostatic activity. The polynucleotide is used to produce a

XX  
CC polypeptide, to detect differentially expressed genes correlated with a

XX  
CC cancerous state of a mammalian cell and to inhibit tumor growth. The

XX  
CC polynucleotide is used as a probe in mapping and tissue profiling. The

XX  
CC encoded polypeptide and antibodies to the polypeptide can also be used

XX  
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for

XX  
CC gene therapy. Note: The sequence data for this patent did not form part

XX  
CC of the printed specification, but was obtained in electronic format

XX  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 550 BP; 140 A; 133 C; 134 G; 143 T; 0 U; 0 Other;

XX  
Query Match 1.2%; Score 25; DB 6; Length 550;

XX  
Best Local Similarity 100.0%; Pred. No. 0.39;

XX  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||  
Db 358 CTTACTAAAAATATATAAAATTAGCTG 382

## RESULT 53

ADN12564/c

ID ADN12564 standard; cDNA; 553 BP.

XX  
AC ADN12564;

XX

DT 29-JUL-2004 (first entry)

XX  
DE Human prostate/colon/lung/breast cancer-related cDNA 79, SEQ.79.

XX  
KW Human; cancer; tumour; prostate cancer; colon cancer; lung cancer;

XX  
KW breast cancer; drug screening; diagnosis; prognosis; prevention;

XX  
KW gene mapping; tissue typing; tissue profiling; cytostatic; gene therapy;

XX  
KW 69.

XX  
OS Homo sapiens.

XX  
PN WO2004039943-A2.

XX  
PD 13-MAY-2004.

XX  
PF 16-MAY-2003; 2003WO-US015465.

XX  
PR 17-MAY-2002; 2002US-0381533P.

XX  
PR 04-FEB-2003; 2003US-0445222P.

XX  
PA (CHIR ) CHIRON CORP.

XX  
PI Scott EM, Lamson G, Kassam A, Zhang G, Sakamoto D, Garcia PD;

XX  
DR MPI; 2004-376173/35.

XX  
PT New isolated polynucleotides, useful for gene mapping or tissue typing or

XX  
PT profiling, as diagnostic reagents, and for preventing or treating cancer,

XX  
PT e.g. prostate, colon, or breast cancer.

XX  
XX Claim 2; SEQ ID NO 79; 190bp; English.

XX  
CC The invention relates to nucleic acids (ADN12486-ADN13970) isolated from

XX  
CC human prostate, colon, lung and breast cancer cDNA libraries, and to 57

XX  
CC proteins (ADN13971-ADN14021) encoded by a subset of these cDNA sequences

XX  
CC (ADN13914-ADN13970). The invention also relates to vectors and host cells

XX  
CC comprising a nucleic acid of the invention; a method for the recombinant

XX  
CC production of a protein of the invention; an antibody specific for a

XX  
CC protein of the invention; a polynucleotide library comprising at least

XX  
CC one nucleic acid sequence of the invention; a method for detecting a

XX  
CC cancerous cell by PCR or probe hybridisation; inhibiting a cancerous

XX  
CC phenotype (particularly aberrant proliferation) of a cell; a method of

XX  
CC identifying an agent that modulates the biological activity of a gene

XX  
CC product differentially expressed in a cancerous cell compared with a

XX  
CC normal cell; and a method of treating a cancer patient using the agent

XX  
CC identified. The nucleic acids and polypeptides can be used to diagnose,

XX  
CC prognose, treat or prevent cancers such as prostate, colon, lung or

XX  
CC breast cancer, and can also be used to screen for drugs for the treatment

XX  
CC of cancer. The nucleic acids can also be used for gene mapping, tissue

XX  
CC typing and tissue profiling. The present sequence represents a

XX  
CC specifically claimed cancer-related cDNA of the invention. Note: The

XX  
CC sequence data for this patent did not form part of the printed

XX  
CC specification, but was obtained in electronic format directly from WIPO

XX  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 553 BP; 140 A; 116 C; 134 G; 163 T; 0 U; 0 Other;

XX  
Query Match 1.2%; Score 25; DB 12; Length 553;

XX  
Best Local Similarity 100.0%; Pred. No. 0.39;

XX  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||  
Db 317 CTTACTAAAAATATATAAAATTAGCTG 293

## RESULT 54

AAH10441/c

ID AAH10441 standard; cDNA; 555 BP.

XX  
AC AAH10441;

XX

DT 26-JUN-2001 (first entry)



```
XX DE Human cDNA clone (3'-primer) SEQ ID NO:7276.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Oca T, Isegai T, Nishikawa T, Hayaeshi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 3; SEQ ID NO 7276; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX SQ Sequence 555 BP; 136 A; 157 C; 136 G; 119 T; 0 U; 7 Other;
XX
XX Query Match 1.2%; Score 25; DB 4; Length 555;
XX Best Local Similarity 100.0%; Pred. No. 0.39;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 167 CTACTAAAAATATATAAATTAGCTG 143

RESULT 55
AAH10596/c
ID AAH10596 standard; cDNA; 555 BP.
XX
XX AC AAH10596;
XX
```

```
DT 26-JUN-2001 (first entry)
XX DE Human cDNA clone (3'-primer) SEQ ID NO:7431.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Oca T, Isegai T, Nishikawa T, Hayaeshi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 3; SEQ ID NO 7431; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX SQ Sequence 555 BP; 144 A; 142 C; 117 G; 150 T; 0 U; 2 Other;
XX
XX Query Match 1.2%; Score 25; DB 4; Length 555;
XX Best Local Similarity 100.0%; Pred. No. 0.39;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 160 CTACTAAAAATATATAAATTAGCTG 136

RESULT 56
ABV51377
ID ABV51377 standard; cDNA; 582 BP.
XX
XX AC ABV51377;
```

```
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 51368.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183119P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 9983-9984; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 582 BP; 182 A; 114 C; 138 G; 146 T; 0 U; 2 Other;
XX
XX Query Match 1.2%; Score 25; DB 5; Length 582;
XX Best Local Similarity 100.0%; Pred. No. 0.39;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2012 CTACTAAAAATATATAAAATTAGCTG 2036
XX |||||||||||||||||||
XX 296 CTACTAAAAATATATAAAATTAGCTG 320
XX
XX RESULT 57
XX ABN60118
XX ID ABN60118 standard; cDNA; 590 BP.
XX
XX AC ABN60118;
XX
XX DT 28-JUN-2002 (first entry)
XX
XX Human cancer related polynucleotide SEQ ID NO 85.
XX
XX Human; cytostatic; gene expression; gene mapping; tissue profiling;
XX gene therapy; cancer; tumour; gene; ss.
XX
XX Homo sapiens.
XX
```

```
XX WO200214500-A2.
XX
XX 21-FEB-2002.
XX
XX 16-AUG-2001; 2001WO-US025840.
XX
XX 16-AUG-2000; 2000US-0226326P.
XX
XX (CHIR ) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F,
XX Lameon G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX
XX WPI; 2002-241905/29.
XX
XX New nucleic acid for producing a polypeptide, detecting differentially
XX expressed genes correlated with a cancerous state of a mammalian cell,
XX and inhibiting tumor growth.
XX
XX Claim 1; SEQ ID NO 85; 883pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
XX with cyostatic activity. The polynucleotide is used to produce a
XX polypeptide, to detect differentially expressed genes correlated with a
XX cancerous state of a mammalian cell and to inhibit tumour growth. The
XX polynucleotide is used as a probe in mapping and tissue profiling. The
XX encoded polypeptide and antibodies to the polypeptide can also be used
XX for therapeutic and diagnostic purposes. The polynucleotide is useful for
XX gene therapy. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 590 BP; 154 A; 151 C; 168 G; 117 T; 0 U; 0 Other;
XX
XX Query Match 1.2%; Score 25; DB 6; Length 590;
XX Best Local Similarity 100.0%; Pred. No. 0.39;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2012 CTACTAAAAATATATAAAATTAGCTG 2036
XX |||||||||||||||||||
XX 418 CTACTAAAAATATATAAAATTAGCTG 442
XX
XX RESULT 58
XX ABV58013
XX ID ABV58013 standard; cDNA; 605 BP.
XX
XX AC ABV58013;
XX
XX DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 58004.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183119P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
```



PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-02444517P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-451937/48.  
XX  
XX Isolated polypeptide for treating, preventing and/or prognosing  
XX PT disorders related to the musculoskeletal system including musculoskeletal  
XX PT cancers and also for testing and detection e.g. diagnosis.  
XX  
XX  
XX Example 2; SEQ ID NO 2901; 781bp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins  
XX CC (ABR03087-ABR04109) associated with the musculoskeletal system useful for  
XX CC preventing, treating or ameliorating medical conditions e.g. by protein  
XX CC or gene therapy. The genes are isolated from a range of human tissues  
XX CC and (ant)agonists are useful in the diagnosis, treatment and prevention  
XX CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the  
XX CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
XX CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,

CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.  
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
CC bacterial, fungal and parasitic infections. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 613 BP; 151 A; 150 C; 129 G; 183 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 1.2%; Score 25; DB 4; Length 613;  
XX Best Local Similarity 100.0%; Pred. No. 0.39;  
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATAAAAATTACCTG 2036  
DB 65 CTACTAAAAATATAAAAATTACCTG 89  
AC  
XX  
XX ABX59524;  
XX  
XX 26-FEB-2003 (first entry)  
XX  
XX  
XX cDNA encoding novel human musculoskeletal system antigen #1868.  
DE  
XX  
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;  
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;  
XX cardiovascular condition; wound; injury; burn; angioneogenesis; ulcer;  
XX post-operative tissue repair; limb regeneration; neuronal growth;  
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
XX AIDS-related complex; chondrocyte growth; bone regeneration;  
XX periodontal regeneration; tissue transport; bone graft; skin aging;  
XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;  
XX cell growth; organ transplant; cell differentiation; body height; weight;  
XX hair colour; eye colour; skin; percentage of adipose tissue;  
XX pigmentation; cosmetic surgery; metabolism; biocrythm; circadian rhythm;  
XX depression; tendency for violence; pain; reproductive capability;  
XX hormone level; endocrine level; appetite; libido; memory; stress;  
XX storage capability; fat content; lipid content; protein content;  
XX carbohydrate content; vitamin content; cofactor content;  
XX nutritional component.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US2002147140-A1.  
XX  
XX 10-OCT-2002.  
XX  
XX  
XX 17-JAN-2001; 2001US-00764877.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 28-JUN-2000; 2000US-0214886P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX PR 07-JUL-2000; 2000US-0216880P.  
XX PR 11-JUL-2000; 2000US-0217487P.  
XX PR 11-JUL-2000; 2000US-0217496P.  
XX PR 14-JUL-2000; 2000US-0218290P.  
XX PR 26-JUL-2000; 2000US-0220863P.  
XX PR 26-JUL-2000; 2000US-0220964P.  
XX PR 14-AUG-2000; 2000US-0224518P.  
XX PR 14-AUG-2000; 2000US-0224519P.  
XX PR 14-AUG-2000; 2000US-0225267P.  
XX PR 14-AUG-2000; 2000US-0225268P.  
XX PR 14-AUG-2000; 2000US-0225270P.  
XX PR 14-AUG-2000; 2000US-0225477P.  
XX PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229333P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
XX (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

DR WPI; 2003-128199/12.

PT Isolated nucleic acid molecules encoding musculoskeletal system  
associated polypeptides, useful for detecting disorders, e.g. cancer.

PS Disclosure; SEQ ID NO 2901; 321bp; English.

XX  
CC The invention describes an isolated nucleic acid molecule comprising a  
CC sequence encoding musculoskeletal system associated polypeptides useful  
CC for detecting disorders, e.g., cancer or cancer metastases, in animals or  
CC humans. The nucleic acid: stimulates re-vascularisation of ischaemic  
CC tissues associated with conditions such as thrombosis, arteriosclerosis,  
CC and other cardiovascular conditions; treats wounds due to injuries,  
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis  
CC and limb regeneration; stimulates neuronal growth; can treat and prevent  
CC neuronal damage occurring in certain disorders or neurodegenerative  
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-  
CC related complex; stimulates chondrocyte growth, thus they can be used to  
CC enhance bone and periodontal regeneration and aid in tissue transports or  
CC bone grafts; prevents skin aging due to sunburn by stimulating  
CC keratinocyte growth; prevents hair loss, since RGF family members  
CC activate hair-forming cells and promotes melanocyte growth; stimulates  
CC growth and differentiation of hematopoietic cells and bone marrow cells  
CC when used in combination with other cytokines; maintains organs before  
CC transplantation or for supporting cell culture of primary tissues;  
CC induces tissue of mesodermal origin to differentiate in early embryos;  
CC increases or decreases the differentiation or proliferation of embryonic  
CC stem cells, besides, haematopoietic lineage; modulates mammalian  
CC characteristics, such as, body height, weight, hair colour, eye colour,  
CC skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,  
CC cosmetic surgery); modulates mammalian metabolism; changes mammal's metal  
CC state or physical state by influencing biorhythms, cardiac rhythms,  
CC depression, tendency for violence, tolerance for pain, reproductive

CC capabilities, hormonal or endocrine levels, appetite, libido, memory, or  
CC stress; increases or decreases storage capabilities, fat content, lipid,  
CC protein, carbohydrate, vitamin, minerals, cofactors or other nutritional  
CC components. This sequence encodes a novel human musculoskeletal system  
CC antigen. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140  
XX

Sequence 613 BP; 151 A; 150 C; 129 G; 183 T; 0 U; 0 Other;  
Query Match 1.2%; Score 25; DB 8; Length 613;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
Db 65 CTACTAAAAATATATAAATTAGCTG 89

RESULT 61  
ADJ30274  
ID ADJ30274 standard; DNA; 613 BP.

XX  
AC ADJ30274;  
XX  
DT 20-MAY-2004 (first entry)

DE Human musculoskeletal system-associated genomic DNA - SEQ ID 2901.

XX  
KM musculoskeletal system; cytoskeletal; osteopathic; cancer; osteoporosis;  
XX gene therapy; vaccine; human; ds.

XX  
XX Homo sapiens.

OS  
PN US2004009488-A1.

XX  
PD 15-JAN-2004.

XX  
PF 13-SEP-2002; 2002US-00242515.

XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-022668P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 01-SEP-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0231968P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240360P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764877.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2004-090458/09.  
XX  
XX New nucleic acid molecule, useful for preparing a medicament for  
PT preventing, treating or ameliorating a medical condition e.g., cancer of  
PT musculoskeletal tissues or osteoporosis.  
XX  
XX  
XX Disclosure; SEQ ID NO 2901; 2899p; English.  
XX  
XX The invention relates to a novel isolated musculoskeletal system-  
CC associated nucleic acid molecule. The nucleic acid of the invention  
CC demonstrates cytostatic and osteopathic activities and may be useful for  
CC preparing a medicament for preventing, treating or ameliorating a medical  
CC condition such as cancer of the musculoskeletal tissues or osteoporosis,  
CC possibly via gene therapy or vaccine production. The current sequence is  
CC that of the human musculoskeletal system-associated genomic DNA of the  
CC invention. The current sequence is not shown within the specification per  
CC se but is available on the USPTO web-site  
CC http://seqdata.uspto.gov/sequence.html?docID=20040009488.  
XX  
XX  
SQ Sequence 613 BP; 151 A; 150 C; 129 G; 183 T; 0 U; 0 Other;  
XX  
XX  
Query Match 1.2%; Score 25; DB 12; Length 613;  
Best Local Similarity 100.0%; Pred. No. 0.39; Mismatches 0; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAATAATTAATAATTAGCTG 2036  
DB 65 CTACTAAATAATTAATAATTAGCTG 89  
RESUL 62  
ADL42725/c  
ID ADL42725 standard; DNA; 684 BP.  
XX  
XX ADL42725;  
AC  
XX

DT 20-MAY-2004 (first entry)  
 XX Human ovarian cancer DNA marker #16615.  
 DE Human; ovarian cancer; de; tumour; cytostatic; DNA marker.  
 XX Homo sapiens.  
 OS WO200170979-A2.  
 PN 27-SEP-2001.  
 XX 21-MAR-2001; 2001WO-US009126.  
 PF 21-MAR-2001; 2000US-0191031P.  
 PR 25-MAY-2000; 2000US-0207124P.  
 PR 15-JUN-2000; 2000US-0211940P.  
 PR 07-JUL-2000; 2000US-0216820P.  
 PR 25-JUL-2000; 2000US-0220661P.  
 PR 21-DEC-2000; 2000US-0257672P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA Lee J, Lillie J;  
 PI WPI; 2001-611502/70.  
 DR Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
 PT cancer cells as compared to their normal non-cancerous ovarian cells are  
 PT used to characterize stage, grade, histological type of ovarian cancer.  
 XX Disclosure; SEQ ID NO 16615; 106pp; English.  
 PS The invention relates to nucleic acid markers which are overexpressed in  
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-  
 CC cancerous) ovarian cells. The invention also relates to polypeptides  
 CC encoded by the markers, antibodies that selectively bind to the  
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
 CC of developing ovarian cancer involving inhibiting expression of a gene  
 CC corresponding to a marker of the invention and a method of treating a  
 CC patient afflicted with ovarian cancer comprising providing to cells of  
 CC the patient an antisense oligonucleotide complementary to a marker of the  
 CC invention. The markers are useful for assessing if a patient is afflicted  
 CC with ovarian cancer, which involves comparing the level of expression of  
 CC a marker in a patient sample and a normal level of expression of the  
 CC marker in a control non-ovarian cancer sample. A difference between the  
 CC expression levels indicates ovarian cancer. The level of expression of a  
 CC marker corresponds to a secreted protein or to a transcribed  
 CC polynucleotide or its portion. The level of expression of the marker is  
 CC assessed by detecting the presence in the sample, a protein or protein  
 CC fragment corresponding to the marker. The presence of protein or protein  
 CC fragment is detected using an antibody that specifically binds with the  
 CC protein or protein fragment. Alternatively, the level of expression of  
 CC the marker is assessed by detecting the presence of a transcribed  
 CC polynucleotide which anneals with the marker or anneals with a portion of  
 CC the polynucleotide comprising the marker, under stringent conditions. The  
 CC marker is also used for monitoring the progression of ovarian cancer in a  
 CC patient which involves detecting expression of the marker in a patient  
 CC sample at a first point in time, repeating the method at a subsequent  
 CC time and comparing the level of expression. The method is carried out  
 CC using an ovarian tissue sample. A composition comprising a marker,  
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
 CC This sequence represents a human ovarian cancer DNA marker of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from Wipo at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 684 BP; 140 A; 157 C; 160 G; 226 T; 0 U; 1 Other;

Query Match 1.2%; Score 25; DB 5; Length 684;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAAATATAAATTAGCTG 2036  
 DB 366 CTAATAAAATATAAATTAGCTG 342

RESULT 63  
 AAD00585  
 ID AAD00585 standard; DNA; 685 BP.  
 XX AAD00585;  
 AC AAD00585;  
 XX 21-SEP-2000 (first entry)  
 DT Alternative version of Human Hscd6 intron 5.  
 DE

XX Hscd6; CSH; human; DNA replication; mitosis; cyclin-dependent kinase;  
 KW cdk; regulator; inhibitor; cdcp; cdc18; vertebrate cdc6; antagonistic;  
 KW agonist; treatment; cell proliferative disease; psoriasis; medicament;  
 KW atherosclerotic vascular disease; vascular restenosis; cardiomyopathy;  
 KW inflammatory arthritis; autoimmune disease; organ transplant rejection;  
 KW traumatic injury; stroke; myocardial infarction; renal failure; antibody;  
 KW hepatic failure; cancer; breast; colon; cervix; lymphoma; diagnosis;  
 KW screen; cytostatic; immunosuppressive; antiinflammatory; antiarthritic;  
 KW antipsoriatic; antiarteriosclerosis; vasotropic; cardiac; vulnary;  
 KW tranquiliser; cerebroprotective; intron; ds.

XX Homo sapiens.  
 OS WO200026242-A2.  
 PN 11-MAY-2000.  
 PD 28-OCT-1999; 99WO-US025445.  
 PF 30-OCT-1998; 98US-00183266.  
 PR (COLD-) COLD SPRING HARBOR LAB.  
 PA (TEXA) UNIV TEXAS SYSTEM.  
 PI Stillman B, Williams RS, Mendez JR;  
 PI WPI; 2000-365574/31.

PT DNA replication-regulating gene, which is vertebrate cdc6 gene and its  
 PT antagonist is useful for treating cell proliferative diseases such as  
 PT atherosclerotic vascular disease, vascular restenosis, psoriasis.

PS Disclosure; Page 100; 11pp; English.

XX The present DNA sequence is an alternative version of intron 5 of the  
 CC human Hscd6 genomic sequence (previously referred as CSH, involved in  
 CC cell cycle regulation. Hscd6 functions as a regulator of DNA replication  
 CC and/or entry of a cell into mitosis and as a cyclin-dependent kinase  
 CC (cdk) inhibitor. It is expressed in active proliferative cells. It has  
 CC similarity to cdc18 and cdc6p proteins from Saccharomyces cerevisiae and  
 CC S. pombe, respectively. The agonist or antagonist of vertebrate cdc6 gene  
 CC is useful for treating cell proliferative diseases like atherosclerotic  
 CC vascular disease, vascular restenosis, psoriasis, inflammatory arthritis,  
 CC autoimmune diseases and organ transplant rejection. Cdc6 gene or agonist  
 CC is useful for treating conditions associated with loss of viable tissues,  
 CC such as traumatic injury, stroke, myocardial infarction, cardiomyopathy,  
 CC renal and hepatic failure, by enhancing cell proliferation. An antagonist  
 CC comprising cdc6 specific antibody is useful to manufacture medicament for  
 CC treatment of cancer of the breast, colon, cervix or lymphoma. The  
 CC antibodies are used to screen for hscd6 and diagnosis of a proliferative  
 CC disorder. Note: This sequence is an alternative version of the human  
 CC Hscd6 intron 5 (Seq ID:NO:16), found in Figure 8D (AAD00997)

XX Sequence 685 BP; 195 A; 118 C; 166 G; 206 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 3; Length 685;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATATATAAATTAGCTG 2036  
Db 303 CTAATAAATATATAAATTAGCTG 327

RESULT 64  
AAL04652  
ID AAL04652 standard; DNA; 692 BP.  
XX  
AC AAL04652;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human reproductive system related antigen DNA SEQ ID NO: 7340.  
XX  
KW Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001339.  
XX  
31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 25-SEP-2000; 2000US-0234999P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239355P.  
PR 13-OCT-2000; 2000US-0239373P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241807P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.



PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249287P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251900P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
XX used in preventing, treating or ameliorating a medical condition.  
XX  
PS Disclosure; SEQ ID NO 7340; 1297bp + Sequence Listing; English.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention  
XX  
SQ Sequence 692 BP; 172 A; 151 C; 187 G; 182 T; 0 U; 0 Other;  
Query Match 1.2%; Score 25; DB 4; Length 692;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2012 CTACTAAATATATAAATTAGCTG 2036  
Db 183 CTACTAAATATATAAATTAGCTG 207  
RESULT 65  
ID AAL04651 standard; DNA; 692 BP.  
XX  
AC AAL04651;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human reproductive system related antigen DNA SEQ ID NO: 7339.  
XX  
KW Human; reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001339.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUL-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0225799P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0228343P.  
PR 01-SEP-2000; 2000US-0228344P.  
PR 01-SEP-2000; 2000US-0228345P.  
PR 05-SEP-2000; 2000US-0228509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233633P.  
PR 14-SEP-2000; 2000US-0233664P.  
PR 14-SEP-2000; 2000US-0233665P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234984P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236803P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.

	PR	20-OCT-2000;	2000US-0240960P.
	PR	20-OCT-2000;	2000US-0241221P.
	PR	20-OCT-2000;	2000US-0241785P.
	PR	20-OCT-2000;	2000US-0241786P.
	PR	20-OCT-2000;	2000US-0241787P.
	PR	20-OCT-2000;	2000US-0241808P.
	PR	20-OCT-2000;	2000US-0241809P.
	PR	20-OCT-2000;	2000US-0241826P.
	PR	01-NOV-2000;	2000US-0244617P.
	PR	08-NOV-2000;	2000US-0246474P.
	PR	08-NOV-2000;	2000US-0246475P.
	PR	08-NOV-2000;	2000US-0246476P.
	PR	08-NOV-2000;	2000US-0246477P.
	PR	08-NOV-2000;	2000US-0246478P.
	PR	08-NOV-2000;	2000US-0246523P.
	PR	08-NOV-2000;	2000US-0246524P.
	PR	08-NOV-2000;	2000US-0246525P.
	PR	08-NOV-2000;	2000US-0246526P.
	PR	08-NOV-2000;	2000US-0246527P.
	PR	08-NOV-2000;	2000US-0246528P.
	PR	08-NOV-2000;	2000US-0246532P.
	PR	08-NOV-2000;	2000US-0246533P.
	PR	08-NOV-2000;	2000US-0246534P.
	PR	08-NOV-2000;	2000US-0246535P.
	PR	08-NOV-2000;	2000US-0246536P.
	PR	08-NOV-2000;	2000US-0246610P.
	PR	08-NOV-2000;	2000US-0246611P.
	PR	08-NOV-2000;	2000US-0246613P.
	PR	17-NOV-2000;	2000US-0249207P.
	PR	17-NOV-2000;	2000US-0249208P.
	PR	17-NOV-2000;	2000US-0249209P.
	PR	17-NOV-2000;	2000US-0249210P.
	PR	17-NOV-2000;	2000US-0249211P.
	PR	17-NOV-2000;	2000US-0249212P.
	PR	17-NOV-2000;	2000US-0249213P.
	PR	17-NOV-2000;	2000US-0249214P.
	PR	17-NOV-2000;	2000US-0249215P.
	PR	17-NOV-2000;	2000US-0249216P.
	PR	17-NOV-2000;	2000US-0249217P.
	PR	17-NOV-2000;	2000US-0249218P.
	PR	17-NOV-2000;	2000US-0249244P.
	PR	17-NOV-2000;	2000US-0249245P.
	PR	17-NOV-2000;	2000US-0249246P.
	PR	17-NOV-2000;	2000US-0249265P.
	PR	17-NOV-2000;	2000US-0249297P.
	PR	17-NOV-2000;	2000US-0249299P.
	PR	17-NOV-2000;	2000US-0249300P.
	PR	01-DEC-2000;	2000US-0250160P.
	PR	01-DEC-2000;	2000US-0250391P.
	PR	05-DEC-2000;	2000US-0251030P.
	PR	05-DEC-2000;	2000US-0251988P.
	PR	05-DEC-2000;	2000US-0256719P.
	PR	06-DEC-2000;	2000US-0251479P.
	PR	08-DEC-2000;	2000US-0251856P.
	PR	08-DEC-2000;	2000US-0251868P.
	PR	08-DEC-2000;	2000US-0251869P.
	PR	08-DEC-2000;	2000US-0251989P.
	PR	08-DEC-2000;	2000US-0251990P.
	PR	11-DEC-2000;	2000US-0254057P.
	PR	05-JAN-2001;	2001US-0259678P.
PX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Rosen CA, Barash SC, Ruben SM,		
XX	WPI; 2001-465570/50 .		
DR			
XX			
PT	Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.		
FT			
XX			
PS	Disclosure; SEQ ID NO 7339; 1297bp + Sequence Listing; English.		
CC	The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a		

Query Match	1.2%: Score 25; DB 4; Length 692;	Best Local Similarity 100.0%; Pred. NO. 0.39;	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
CC	Protein of the invention		
XX			
SQL	Sequence 692 BP; 172 A; 151 C; 187 G; 182 T; 0 U; 0 Other;		
QY	2012 CTACTAAATAATATAAATTAGCTG 2036		
DB	183 CTACTAAATAATATAAATTAGCTG 207		
RESULT 66			
AAH92601/c			
ID	AAH92601 standard; DNA; 700 BP.		
XX			
AC	AAH92601;		
XX			
DT	09-OCT-2001 (first entry)		
XX			
DE	Human inflammatory bowel disease related gene fragment ICR1301a.		
XX			
KM	Human; inflammatory bowel disease; Crohn's disease; ulcerative colitis;		
KW	single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;		
KX	chromosome 5q31-33; forensic test; gene therapy; ds.		
XX			
OS	Homo sapiens.		
PN	WO200142511-A2.		
PD	14-JUN-2001.		
XX			
PF	11-DEC-2000; 2000WO-US033632.		
XX			
PR	10-DEC-1999; 99US-0170257P.		
PR	10-APR-2000; 2000US-0196046P.		
XX			
PA	(WHED ) WHITEHEAD INST BIOMEDICAL RES.		
PA	(ELLI-) ELLIPSIS BIOTHERAPEUTICS CORP.		
XX			
PI	Daly M, Hudson TJ, Lander ES, Rioux J, Siminovitch K;		
DR	WPI; 2001-367874/38.		
XX			
PT	Testing for the presence of polymorphisms associated with inflammatory		
PT	bowel disease, using a hybridization assay.		
XX			
PS	Disclosure; Page 264; 463pp; English.		
XX			
CC	The present invention describes a method for detecting the presence of		
CC	polymorphisms associated with inflammatory bowel diseases such as		
CC	ulcerative colitis and Crohn's disease. The methods can be used to detect		
CC	the presence of genetic polymorphisms associated with inflammatory bowel		
CC	disease and correlating their occurrence with disease states. They may be		
CC	used in this way for phenotypic correlations, forensics, paternity		
CC	testing, medicine and genetic analysis. The present sequence is a gene		
CC	containing a polymorphic site described in the exemplification of the		
CC	invention		
XX			
SQL	Sequence 700 BP; 142 A; 174 C; 137 G; 247 T; 0 U; 0 Other;		
QY	Query Match 1.2%; Score 25; DB 4; Length 700;		
DB	Best Local Similarity 100.0%; Pred. NO. 0.39;		
XX	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
QY	2012 CTACTAAATAATATAAATTAGCTG 2036		
DB	534 CTACTAAATAATATAAATTAGCTG 510		
RESULT 67			
AAH92602/c			

ID AAH92602 standard; DNA; 700 BP.  
XX  
AC AAH92602;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Human inflammatory bowel disease related gene fragment IGRI302a.  
XX  
KM Human; inflammatory bowel disease; Crohn's disease; ulcerative colitis;  
KW single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;  
XX chromosome 5q31-33; forensic test; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200142511-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 11-DEC-2000; 2000MO-US033632.  
XX  
PR 10-DEC-1999; 99US-0170257P.  
XX  
PR 10-APR-2000; 2000US-0196046P.  
XX  
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
XX (ELI-) ELIIPSIS BIOTHERAPEUTICS CORP.  
XX  
PI Daly M, Hudson TJ, Lander ES, Rioux J, Siminovitch K;  
XX  
DR WPI; 2001-367874/38.  
XX  
PT Testing for the presence of polymorphisms associated with inflammatory  
XX bowel disease, using a hybridization assay.  
XX  
PS Disclosure; Page 264-265; 463pp; English.  
XX  
CC The present invention describes a method for detecting the presence of  
CC polymorphisms associated with inflammatory bowel diseases such as  
CC ulcerative colitis and Crohn's disease. The methods can be used to detect  
CC the presence of genetic polymorphisms associated with inflammatory bowel  
CC disease and correlating their occurrence with disease states. They may be  
CC used in this way for phenotypic correlations, forensics, paternity  
CC testing, medicine and genetic analysis. The present sequence is a gene  
CC containing a polymorphic site described in the exemplification of the  
CC invention  
XX  
SQ Sequence 700 BP; 148 A; 186 C; 143 G; 223 T; 0 U; 0 Other;  
XX  
Query Match 1.2%; Score 25; DB 4; Length 700;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 35 CTACTAAAAATATATAAATTAGCTG 11  
XX  
RESULT 68  
ABV42610  
ID ABV42610 standard; cDNA; 703 BP.  
XX  
AC ABV42610;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 42601.  
XX  
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX

PD 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
PF 17-FEB-2000; 2000US-0183319P.  
XX  
PR 16-MAR-2000; 2000US-0189862P.  
XX  
PR 25-MAY-2000; 2000US-0207454P.  
XX  
PR 09-JUN-2000; 2000US-0211314P.  
XX  
PR 18-JUL-2000; 2000US-0219007P.  
XX  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JB;  
XX  
DR WPI; 2001-662795/76.  
XX  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 8525-8526; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 703 BP; 223 A; 141 C; 148 G; 190 T; 0 U; 1 Other;  
XX  
Query Match 1.2%; Score 25; DB 5; Length 703;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 228 CTACTAAAAATATATAAATTAGCTG 252  
XX  
RESULT 69  
ADL62169/C  
ID ADL62169 standard; DNA; 742 BP.  
XX  
AC ADL62169;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human ovarian cancer DNA marker #20381.  
XX  
KM Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.  
KW Homo sapiens.  
XX  
OS Homo sapiens.  
XX  
PN WO200170979-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009126.  
XX  
PR 21-MAR-2000; 2000US-0191031P.  
XX  
PR 25-MAY-2000; 2000US-0207124P.  
XX  
PR 15-JUN-2000; 2000US-0211940P.  
XX  
PR 07-JUL-2000; 2000US-0216820P.  
XX  
PR 25-JUL-2000; 2000US-0220661P.  
XX  
PR 21-DEC-2000; 2000US-0257672P.  
XX

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA Lee J, Lillie J;  
 PI WPI; 2001-611502/70.  
 DR  
 XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
 PT cancer cells as compared to their normal non-cancerous ovarian cells are  
 PT used to characterize stage, grade, histological type of ovarian cancer.  
 XX  
 PS Disclosure; SEQ ID NO 20381; 106bp; English.  
 XX  
 XX The invention relates to nucleic acid markers which are overexpressed in  
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-  
 CC cancerous) ovarian cells. The invention also relates to polypeptides  
 CC encoded by the markers, antibodies that selectively bind to the  
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
 CC of developing ovarian cancer involving inhibiting expression of a gene  
 CC corresponding to a marker of the invention and a method of treating a  
 CC patient afflicted with ovarian cancer comprising providing to cells of  
 CC the patient an antisense oligonucleotide complementary to a marker of the  
 CC invention. The markers are useful for assessing if a patient is afflicted  
 CC with ovarian cancer, which involves comparing the level of expression of  
 CC a marker in a patient sample and a normal level of expression of the  
 CC marker in a control non-ovarian cancer sample. A difference between the  
 CC expression levels indicates ovarian cancer. The level of expression of a  
 CC marker corresponds to a secreted protein or to a transcribed  
 CC polynucleotide or its portion. The level of expression of the marker is  
 CC assessed by detecting the presence in the sample, a protein or protein  
 CC fragment corresponding to the marker. The presence of protein or protein  
 CC fragment is detected using an antibody that specifically binds with the  
 CC protein or protein fragment. Alternatively, the level of expression of  
 CC the marker is assessed by detecting the presence of a transcribed  
 CC polynucleotide which anneals with the marker or anneals with a portion of  
 CC the polynucleotide comprising the marker, under stringent conditions. The  
 CC marker is also used for monitoring the progression of ovarian cancer in a  
 CC patient which involves detecting expression of the marker in a patient  
 CC sample at a first point in time, repeating the method at a subsequent  
 CC time and comparing the level of expression. The method is carried out  
 CC using an ovarian tissue sample. A composition comprising a marker,  
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
 CC This sequence represents a human ovarian cancer DNA marker of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 742 BP; 172 A; 207 C; 156 G; 204 T; 0 U; 3 Other;  
 SQ  
 Query Match 1.2%; Score 25; DB 5; Length 742;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
 DB 567 CTTACTAAAAATATATAAATTAGCTG 543  
 RESULT 70  
 ID ADB82582/c  
 ID ADB82582 standard; cDNA; 743 BP.  
 XX  
 AC ADB82582;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 XX Human cDNA sequence useful for the treatment of cancer (SegID 894).  
 DE  
 XX  
 XX human; prostate; cancer; cytostatic; gene therapy; vaccine;  
 KM immune response; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX

PN WO2003050236-A2.  
 XX  
 PD 19-JUN-2003.  
 XX  
 XX 04-SEP-2002; 2002WO-US0282214.  
 PF  
 PR 07-DEC-2001; 2001US-00012697.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (HYSR-) HYSR INC.  
 XX  
 PI Escobedo J, Garcia PD, Kassam A, Lamsan G, Drmanac R;  
 PI Cirkvajakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;  
 PI Garcia V, Jones LW, Stache-Crain B, Scott EM;  
 XX  
 DR WPI; 2003-513972/48.  
 XX  
 PT New polynucleotides derived from human prostate, useful for modulating  
 PT immune response to prevent or treat cancer.  
 XX  
 PS Claim 1; SEQ ID NO 894; 188bp; English.  
 XX  
 XX This invention relates to novel isolated polynucleotides of human origin,  
 CC particularly isolated from the human prostate. Specifically, it refers to  
 CC the diagnostics and therapeutics comprising these novel human  
 CC polynucleotides, and includes the derived probes, antisense  
 CC oligonucleotides and antibodies thereof. The identification of these  
 CC human prostate genes that can inhibit tumour growth is useful for  
 CC understanding the progression and nature of complex diseases such as  
 CC cancer, and hence they are important in the drug discovery process. The  
 CC present invention describes these polynucleotides and encoded  
 CC polypeptides as exhibiting cytostatic activity, and through gene therapy  
 CC and/or vaccines they can be used to modulate the immune response for the  
 CC prevention or treatment of cancers, particularly of the prostate, but  
 CC also for breast, lung and colon cancer. This polynucleotide sequence is a  
 CC human cDNA sequence useful for the treatment of cancer, used in an  
 CC exemplification of the invention. NOTE: These sequences are not given in  
 CC the specification but are provided on the WIPO website.  
 CC  
 XX Sequence 743 BP; 169 A; 161 C; 177 G; 224 T; 0 U; 12 Other;  
 SQ  
 Query Match 1.2%; Score 25; DB 9; Length 743;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
 DB 229 CTTACTAAAAATATATAAATTAGCTG 205  
 RESULT 71  
 ID AAL24632/c  
 ID AAL24632 standard; cDNA; 749 BP.  
 XX  
 AC AAL24632;  
 XX  
 DT 07-DEC-2001 (first entry)  
 XX  
 XX Human breast cancer expressed polynucleotide 17089.  
 DE  
 XX  
 XX Human; breast cancer; cell marker; cytostatic; ss.  
 KM  
 OS Homo sapiens.  
 XX  
 PN WO200151628-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 10-JAN-2001; 2001WO-US000798.  
 PD  
 XX  
 PR 14-JAN-2000; 2000US-0176077P.  
 PR 14-MAR-2000; 2000US-0189167P.  
 PR 24-MAR-2000; 2000US-0192099P.  
 PR

PR 29-MAR-2000; 2000US-0193480P.  
 PR 15-MAY-2000; 2000US-0205230P.  
 PR 09-JUN-2000; 2000US-0211315P.  
 PR 25-JUL-2000; 2000US-0220534P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI LiJie J, Xu Y, Wang Y, Steinmann K;  
 XX  
 DR WPI; 2001-451856/48.  
 XX  
 PT New peptide useful as a marker for the diagnosis of breast cancer.  
 XX  
 PS Claim 1; Page 3143-3144; 3695pp; English.  
 XX  
 CC The invention relates to human breast cancer expressed polynucleotides  
 CC (AA07544-AA26789) and method of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterizing treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity  
 CC  
 SQ Sequence 749 BP; 131 A; 221 C; 200 G; 197 T; 0 U; 0 Other;  
 XX  
 Query Match 1.2%; Score 25; DB 4; Length 749;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTACTAAATAATTAATAATTAGCTG 2036  
 DB 178 CTACTAAATAATTAATAATTAGCTG 154  
 XX  
 RESULT 72  
 AA195858/c  
 ID AA195858 standard; cDNA; 781 BP.  
 XX  
 AC AA195858;  
 XX  
 DT 13-NOV-2001 (first entry)  
 XX  
 DE Human neuroblastoma expressed polynucleotide SEQ ID NO 1933.  
 XX  
 KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200166719-A1.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 02-MAR-2001; 2001WO-JP001629.  
 XX  
 PR 07-MAR-2000; 2000JP-00159195.  
 XX  
 PA (CHIB-) CHIBA PREFECTURE.  
 XX  
 PA (HISM) HISAMITSU PHARM CO LTD.  
 XX  
 PI Nakagawara A;  
 XX  
 DR WPI; 2001-565584/63.  
 XX  
 PT Nucleic acids originating in gene expressed in human neuroblastoma,  
 PT useful as probe or primer in diagnosing prognosis of human neuroblastoma,  
 PT malignancy and susceptibility indicator or tumor marker for anti-cancer  
 PT agents.  
 XX  
 PS Claim 1; Page 1430; 2979pp; Japanese.  
 XX  
 CC The invention relates to novel genes (AA193926-AA197963) expressed in

CC human neuroblastoma. The nucleic acids are applicable as a probe or  
 CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and  
 CC susceptibility indicators or tumor markers for anti-cancer agents. The  
 CC gene information for diagnosing prognosis is related to factors similar  
 CC to that for N-myc and TrkA genes  
 CC  
 SQ Sequence 781 BP; 217 A; 174 C; 154 G; 216 T; 0 U; 20 Other;  
 XX  
 Query Match 1.2%; Score 25; DB 4; Length 781;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2012 CTACTAAATAATTAATAATTAGCTG 2036  
 DB 207 CTACTAAATAATTAATAATTAGCTG 183  
 XX  
 RESULT 73  
 AB089641/c  
 ID AB089641 standard; cDNA; 816 BP.  
 XX  
 AC AB089641;  
 XX  
 DT 27-SEP-2002 (first entry)  
 XX  
 DE Human prostate expressed polynucleotide SEQ ID NO 897.  
 XX  
 KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200255700-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 07-DEC-2001; 2001WO-US047349.  
 XX  
 PR 07-DEC-2000; 2000US-0254648P.  
 XX  
 PR 13-MAR-2001; 2001US-0275688P.  
 XX  
 PA (CHIR) CHIRON CORP.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R,  
 PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D,  
 PI Garcia V, Jones WL, Steche-Crain B, Scott EM;  
 XX  
 DR WPI; 2002-557824/59.  
 XX  
 PT New genes and gene products isolated from human prostate, useful for  
 PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast  
 PT cancer), or as vaccines for treating or preventing these diseases.  
 XX  
 PS Claim 1; SEQ ID NO 897; 186pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated polynucleotide comprising any of  
 CC 1477 sequences or its fragment, degenerate variant, antisense or  
 CC complement. The polynucleotides and gene products are useful for treating  
 CC or diagnosing tumor or cancer (e.g. prostate cancer, breast cancer, lung  
 CC cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,  
 CC rabbits, horse or human). The polynucleotides and polypeptides are also  
 CC useful as vaccines for treating or preventing these diseases. The  
 CC polynucleotides are useful for gene therapy. The present sequence is that  
 CC of one of a group of polynucleotides (AB088745-AB090015) disclosed  
 CC electronically as sequences of the invention. However only 1271  
 CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91  
 CC proteins are claimed. Note: The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequence  
 XX  
 SQ Sequence 816 BP; 182 A; 176 C; 194 G; 244 T; 0 U; 20 Other;

Query Match 1.2%; Score 25; DB 6; Length 816;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAAATATATAAATTAGCTG 2036  
 |||||  
 DB 302 CTAATAAAATATATAAATTAGCTG 278

RESULT 74  
 ABX92120/C  
 ID ABX92120 standard; cDNA; 852 BP.  
 XX  
 AC ABX92120;  
 XX  
 DT 08-MAY-2003 (first entry)  
 XX  
 DE Lung specific nucleic acid (LSNA) #162.  
 XX  
 XX Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;  
 KM cancer monitoring; cancer staging; cancer imaging; lung cancer;  
 KM non-cancerous diseases of the lung; transgenic animal; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200268633-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-NOV-2001; 2001WO-US043612.  
 XX  
 PR 22-NOV-2000; 2000US-0252500P.  
 XX  
 PA (DIAD-) DIADEXUS INC.  
 XX  
 PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;  
 DR WPI; 2002-713376/77.  
 XX  
 PT New isolated human nucleic acid molecule and polypeptide, useful for  
 PT identifying, diagnosing, monitoring, staging, imaging and treating lung  
 PT cancer and non-cancerous diseases of the lung.  
 XX  
 PS Claim 1; Page 308; 389pp; English.  
 XX  
 CC The invention describes an isolated human nucleic acid (I) encoding any  
 CC of 120 10-1533 residue amino acid sequences (S1), given in the  
 CC specification, comprising any of 164 179-12421 base pair sequences (S2),  
 CC given in the specification. The methods and compositions of the present  
 CC invention are useful for identifying, diagnosing, monitoring, staging,  
 CC imaging and treating lung cancer and non-cancerous diseases of the lung.  
 CC They are also used for identifying lung tissue, monitoring and  
 CC identifying and/or designing antagonists of the polypeptide of the  
 CC invention, gene therapy, production of transgenic animals and production  
 CC of engineered lung tissue for treatment and research. This sequence  
 CC encodes a lung specific nucleic acid  
 XX  
 SQ Sequence 852 BP; 329 A; 188 C; 174 G; 161 T; 0 U; 0 Other;  
 XX

Query Match 1.2%; Score 25; DB 6; Length 852;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAAATATATAAATTAGCTG 2036  
 |||||  
 DB 53 CTAATAAAATATATAAATTAGCTG 29

RESULT 75  
 AAC57995/C  
 ID AAC57995 standard; DNA; 1001 BP.  
 XX  
 AC AAC57995;  
 XX

XX  
 XX 25-JAN-2001 (first entry)  
 XX  
 DE Arachidonic acid metabolism related genomic biallelic marker #629.  
 XX  
 KM Human; biallelic marker; arachidonic acid metabolism; genotyping;  
 KM detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;  
 KM single nucleotide polymorphism; hybridisation assay; sequencing assay;  
 KM specific amplification assay; identification; ERBM; 12-LO-ERM;  
 KM eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200047771-A2.  
 XX  
 PD 17-AUG-2000.  
 XX  
 PF 11-FEB-2000; 2000WO-1B000184.  
 XX  
 PR 12-FEB-1999; 99US-0119917P.  
 PR 23-MAR-1999; 99US-00275267.  
 PR 07-MAY-1999; 99US-0133200P.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Blumenfeld M, Bougueleret L, Chumakov I;  
 DR WPI; 2000-571681/53.  
 XX  
 PT Novel biallelic markers useful for detecting conditions and genotypes  
 PT associated with arachidonic acid metabolism.  
 XX  
 PS Claim 13; Page 753; 802pp; English.  
 XX  
 CC The present invention describes polynucleotides including biallelic  
 CC markers derived from genes involved in arachidonic acid metabolism and  
 CC from genomic regions flanking those genes. Methods from the present  
 CC invention may be used to select individuals for clinical trials and  
 CC predict responses to treatment with drugs. The polynucleotides may be  
 CC used in hybridisation assays, sequencing assays and specific  
 CC amplification assays for identifying an eicosanoid-related biallelic  
 CC marker (ERBM) or 12-LO-related biallelic marker, and for amplifying a  
 CC segment of nucleotides containing an ERBM. The polynucleotides are useful  
 CC in diagnostic kits. The markers may be used to detect conditions and  
 CC genotypes associated with arachidonic acid metabolism. AAC57367 to  
 CC CAC58019 and AAB24019 and AAB24020 represent sequences used in the  
 CC exemplification of the present invention. N.B. Polymorphic bases (single  
 CC nucleotide polymorphisms also known as SNPs) in the polynucleotide  
 CC sequences from the present invention have been given as their  
 CC corresponding degenerate bases e.g. a polymorphic base of C or T has been  
 CC given as Y  
 XX  
 SQ Sequence 1001 BP; 230 A; 283 C; 233 G; 254 T; 0 U; 1 Other;  
 XX

Query Match 1.2%; Score 25; DB 3; Length 1001;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAAATATATAAATTAGCTG 2036  
 |||||  
 DB 733 CTAATAAAATATATAAATTAGCTG 709

RESULT 76  
 AAC57996/C  
 ID AAC57996 standard; DNA; 1001 BP.  
 XX  
 AC AAC57996;  
 XX  
 DT 25-JAN-2001 (first entry)  
 XX  
 DE Arachidonic acid metabolism related genomic biallelic marker #630.  
 XX

KM Human; biallelic marker; arachidonic acid metabolism; genotyping;  
 KM detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;  
 KM single nucleotide polymorphism; hybridisation assay; sequencing assay;  
 KM specific amplification assay; identification; ERBM; 12-LO-RBM;  
 KM eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200047771-A2.  
 XX  
 PD 17-AUG-2000.  
 XX  
 PF 11-FEB-2000; 2000MO-IB000184.  
 XX  
 PR 12-FEB-1999; 99US-0119917P.  
 PR 23-MAR-1999; 99US-00275267.  
 PR 07-MAY-1999; 99US-0133200P.  
 XX  
 PA (GESP) GENSET.  
 PI Blumenfeld M, Bougueleret L, Chumakov I;  
 DR WPI; 2000-571881/53.  
 XX  
 PT Novel biallelic markers useful for detecting conditions and genotypes  
 PT associated with arachidonic acid metabolism.  
 PS Claim 13; Page 753-754; 802pp; English.  
 XX  
 CC The present invention describes polynucleotides including biallelic  
 CC markers derived from genes involved in arachidonic acid metabolism and  
 CC from genomic regions flanking those genes. Methods from the present  
 CC invention may be used to select individuals for clinical trials and  
 CC predict responses to treatment with drugs. The polynucleotides may be  
 CC used in hybridisation assays, sequencing assays and specific  
 CC amplification assays for identifying an eicosanoid-related biallelic  
 CC marker (ERBM) or 12-LO-related biallelic marker, and for amplifying a  
 CC segment of nucleotides containing an ERBM. The polynucleotides are useful  
 CC in diagnostic kits. The markers may be used to detect conditions and  
 CC genotypes associated with arachidonic acid metabolism. AACS7367 to  
 CC AAC56018 and AAB24019 and AAB24020 represent sequences used in the  
 CC exemplification of the present invention. N.B. Polymorphic bases (single  
 CC nucleotide polymorphisms also known as SNPs) in the polynucleotide  
 CC sequences from the present invention have been given as their  
 CC corresponding degenerate bases e.g. a polymorphic base of C or T has been  
 CC given as Y  
 CC  
 SQ Sequence 1001 BP; 276 A; 253 C; 229 G; 242 T; 0 U; 1 Other;  
 XX  
 Query Match 1.2%; Score 25; DB 3; Length 1001;  
 Best Local Similarity 100.0%; Pred.No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2012 CTAATAAATATATAAATTAGCTG 2036  
 Db 991 CTAATAAATATATAAATTAGCTG 967  
 XX  
 RESULT 77  
 AAH94439  
 ID AAH94439 standard; cDNA; 1055 BP.  
 XX  
 AC AAH94439;  
 XX  
 DT 05-OCT-2001 (first entry)  
 XX  
 DE Human foetal cDNA, SEQ ID NO: 1126.  
 XX  
 KM Human; foetal protein; cytosolic; immunosuppressive; immunostimulant;  
 KM nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;  
 KM gene therapy; antisense therapy; cancer; immune disorder;  
 KM growth disorder; osteoporosis; thrombolytic disorder;  
 KM nervous system disorder; inflammation; expressed sequence tag; EST; ss.

XX  
 OS Homo sapiens.  
 XX  
 PN MO200155339-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001MO-US002723.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 PR 15-SEP-2000; 2000US-00663870.  
 PR 06-NOV-2000; 2000US-00707351.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dorman RA, Tang YT;  
 PI Liu C, Asundi V, Zhou P, Wehrman T;  
 DR WPI; 2001-465571/50.  
 DR P-PSDB; AAM06764.  
 XX  
 PT Novel fetal proteins useful for the treatment and diagnosis of diseases  
 PT associated with dysfunction of the protein e.g. cancers, immune  
 PT disorders, growth disorders, thrombolytic disorders, nervous system  
 PT disorders and inflammation.  
 PS Example 3; Page 630-631; 715pp; English.  
 XX  
 CC The invention relates to novel foetal polypeptides encoded by  
 CC polynucleotides comprising one of 477 sequences fully defined in the  
 CC specification. The foetal polynucleotides and polypeptides are useful in  
 CC the treatment and diagnosis of diseases such as cancers, immune  
 CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,  
 CC nervous system disorders and inflammation. The present sequence is a full  
 CC length cDNA which was assembled using expressed sequence tags (ESTs)  
 CC found to be expressed in human foetal tissue cDNA libraries as seeds  
 XX  
 SQ Sequence 1055 BP; 286 A; 243 C; 269 G; 257 T; 0 U; 0 Other;  
 XX  
 Query Match 1.2%; Score 25; DB 5; Length 1055;  
 Best Local Similarity 100.0%; Pred.No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2012 CTAATAAATATATAAATTAGCTG 2036  
 Db 879 CTAATAAATATATAAATTAGCTG 903  
 XX  
 RESULT 78  
 AAH94410  
 ID AAH94410 standard; cDNA; 1055 BP.  
 XX  
 AC AAH94410;  
 XX  
 DT 05-OCT-2001 (first entry)  
 XX  
 DE Human foetal cDNA, SEQ ID NO: 939.  
 XX  
 KM Human; foetal protein; cytosolic; immunosuppressive; immunostimulant;  
 KM nootropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;  
 KM gene therapy; antisense therapy; cancer; immune disorder;  
 KM growth disorder; osteoporosis; thrombolytic disorder;  
 KM nervous system disorder; inflammation; expressed sequence tag; EST; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200155339-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001MO-US002723.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.

```
PR 15-SEP-2000; 2000US-00663870.
PR 06-NOV-2000; 2000US-00707351.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
PI Liu C, Asundi V, Zhou P, Werhman T;
XX
DR WPI; 2001-465571/50.
DR P-PSDB; AAM06735.
XX
XX Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune
PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation.
XX
XX Claim 1; Page 549; 715pp; English.
XX
XX The invention relates to novel foetal polypeptides encoded by
CC polynucleotides comprising one of 477 sequences fully defined in the
CC specification. The foetal polynucleotides and polypeptides are useful in
CC the treatment and diagnosis of diseases such as cancers, immune
CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
CC nervous system disorders and inflammation. The present sequence is a full
CC length cDNA which was assembled using expressed sequence tags (ESTs)
CC found to be expressed in human foetal tissue cDNA libraries as seeds
XX
SQ Sequence 1055 BP; 286 A; 243 C; 269 G; 257 T; 0 U; 0 Other;

Query Match      1.2%; Score 25; DB 5; Length 1055;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATAAAAATTAGCTG 2036
      |||||||
DB      879 CTACTAAAAATATAAAAATTAGCTG 903

RESULT 79
ABV28056/c
ID ABV28056 standard; cDNA; 1107 BP.
XX
XX ABV28056;
AC
XX
XX 16-SEP-2002 (first entry)
DT
XX
DE Human prostate expression marker cDNA 28047.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX 17-FEB-2000; 2000US-0183319P.
PR
XX 16-MAR-2000; 2000US-0189862P.
PR
XX 25-MAY-2000; 2000US-0207454P.
PR
XX 09-JUN-2000; 2000US-0211314P.
PR
XX 18-JUL-2000; 2000US-0219007P.
PR
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
XX WPI; 2001-662795/76.
DR
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
```

```
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5798; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1107 BP; 329 A; 228 C; 197 G; 335 T; 0 U; 18 Other;

Query Match      1.2%; Score 25; DB 5; Length 1107;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATAAAAATTAGCTG 2036
      |||||||
DB      880 CTACTAAAAATATAAAAATTAGCTG 856

RESULT 80
AAL03604/c
ID AAL03604 standard; DNA; 1229 BP.
XX
XX AAL03604;
AC
XX
XX 21-NOV-2001 (first entry)
DT
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 6292.
XX
XX Human; reproductive system related antigen; reproductive system disorder;
KM cancer; gene therapy; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200155320-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001339.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR
XX 04-FEB-2000; 2000US-0180628P.
PR
XX 24-FEB-2000; 2000US-0184664P.
PR
XX 02-MAR-2000; 2000US-0186350P.
PR
XX 16-MAR-2000; 2000US-0189874P.
PR
XX 17-MAR-2000; 2000US-0190076P.
PR
XX 18-APR-2000; 2000US-0198123P.
PR
XX 19-MAY-2000; 2000US-0205515P.
PR
XX 07-JUN-2000; 2000US-0209467P.
PR
XX 28-JUN-2000; 2000US-0214886P.
PR
XX 30-JUN-2000; 2000US-0215135P.
PR
XX 07-JUL-2000; 2000US-0216647P.
PR
XX 07-JUL-2000; 2000US-0216880P.
PR
XX 11-JUL-2000; 2000US-0217487P.
PR
XX 11-JUL-2000; 2000US-0217496P.
PR
XX 14-JUL-2000; 2000US-0218290P.
PR
XX 26-JUL-2000; 2000US-0220963P.
PR
XX 26-JUL-2000; 2000US-0220964P.
PR
XX 14-AUG-2000; 2000US-0224518P.
PR
XX 14-AUG-2000; 2000US-0224519P.
PR
XX 14-AUG-2000; 2000US-0225213P.
PR
XX 14-AUG-2000; 2000US-0225214P.
PR
XX 14-AUG-2000; 2000US-0225266P.
```



[illegible]

AC ABA07805;  
XX  
DT 11-JUN-2002 (first entry)  
XX  
DE Human ovarian and breast cancer associated polynucleotide SEQ ID NO 600.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; ds.  
OS Homo sapiens.  
XX  
PN MO200155325-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001345.  
XX  
PR 31-JUN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189877P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JUN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-488786/53.  
XX  
XX  
XX New isolated ovarian and/or breast cancer related nucleic acids and  
PT polypeptides, useful for diagnosing, treating and/or preventing human  
PT diseases and disorders, particularly ovarian and/or breast cancer.  
XX  
XX Disclosure, SEQ ID NO 600; 577bp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABA07454-ABA08224) and proteins  
CC (AB010743-AB010980) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune hemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infections diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 1229 BP; 324 A; 252 C; 230 G; 423 T; 0 U; 0 Other;  
XX  
XX Query Match 1.2%; Score 25; DB 4; Length 1229;  
XX Beel Local Similarity 100.0%; Pred. No. 0.38;  
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 2012 CTACTAAATATATATATATAGCTG 2036  
XX |||||  
XX 311 CTACTAAATATATATATATAGCTG 287  
XX  
XX RESULT 82  
XX ID AAK81121 standard; DNA; 1255 BP.  
XX AAK81121;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35933.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200157182-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
XX

PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-022513P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-022557P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-022579P.  
PR 22-AUG-2000; 2000US-022681P.  
PR 22-AUG-2000; 2000US-022686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233297P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0234240P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.

XX	02-OCT-2000,	2000US-0237037P.
PR	02-OCT-2000,	2000US-0237038P.
PR	02-OCT-2000,	2000US-0237039P.
PR	02-OCT-2000,	2000US-0237040P.
PR	13-OCT-2000,	2000US-0239935P.
PR	13-OCT-2000,	2000US-0239937P.
PR	20-OCT-2000,	2000US-0240960P.
PR	20-OCT-2000,	2000US-0241221P.
PR	20-OCT-2000,	2000US-0241785P.
PR	20-OCT-2000,	2000US-0241786P.
PR	20-OCT-2000,	2000US-0241787P.
PR	20-OCT-2000,	2000US-0241808P.
PR	20-OCT-2000,	2000US-0241809P.
PR	20-OCT-2000,	2000US-0241826P.
PR	01-NOV-2000,	2000US-0244617P.
PR	08-NOV-2000,	2000US-0246474P.
PR	08-NOV-2000,	2000US-0246475P.
PR	08-NOV-2000,	2000US-0246476P.
PR	08-NOV-2000,	2000US-0246477P.
PR	08-NOV-2000,	2000US-0246478P.
PR	08-NOV-2000,	2000US-0246523P.
PR	08-NOV-2000,	2000US-0246524P.
PR	08-NOV-2000,	2000US-0246525P.
PR	08-NOV-2000,	2000US-0246526P.
PR	08-NOV-2000,	2000US-0246527P.
PR	08-NOV-2000,	2000US-0246528P.
PR	08-NOV-2000,	2000US-0246532P.
PR	08-NOV-2000,	2000US-0246609P.
PR	08-NOV-2000,	2000US-0246610P.
PR	08-NOV-2000,	2000US-0246611P.
PR	17-NOV-2000,	2000US-0249207P.
PR	17-NOV-2000,	2000US-0249208P.
PR	17-NOV-2000,	2000US-0249209P.
PR	17-NOV-2000,	2000US-0249210P.
PR	17-NOV-2000,	2000US-0249211P.
PR	17-NOV-2000,	2000US-0249212P.
PR	17-NOV-2000,	2000US-0249213P.
PR	17-NOV-2000,	2000US-0249214P.
PR	17-NOV-2000,	2000US-0249215P.
PR	17-NOV-2000,	2000US-0249216P.
PR	17-NOV-2000,	2000US-0249217P.
PR	17-NOV-2000,	2000US-0249218P.
PR	17-NOV-2000,	2000US-0249244P.
PR	17-NOV-2000,	2000US-0249245P.
PR	17-NOV-2000,	2000US-0249246P.
PR	17-NOV-2000,	2000US-0249247P.
PR	17-NOV-2000,	2000US-0249289P.
PR	17-NOV-2000,	2000US-0249289P.
PR	17-NOV-2000,	2000US-0249300P.
PR	01-DEC-2000,	2000US-0250160P.
PR	01-DEC-2000,	2000US-0250391P.
PR	05-DEC-2000,	2000US-0251030P.
PR	05-DEC-2000,	2000US-0251988P.
PR	05-DEC-2000,	2000US-0256719P.
PR	06-DEC-2000,	2000US-0251479P.
PR	08-DEC-2000,	2000US-0251856P.
PR	08-DEC-2000,	2000US-0251868P.
PR	08-DEC-2000,	2000US-0251869P.
PR	08-DEC-2000,	2000US-0251989P.
PR	08-DEC-2000,	2000US-0251990P.
PR	11-DEC-2000,	2000US-0254097P.
PR	05-JAN-2001,	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
XX	WPI; 2001-483426/52.	
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
XX	useful for preventing, diagnosing and/or treating cancers and metastasis	

PS	Disclosure: SEQ ID NO 35933; 3071bp + Sequence Listing; English.
XX	
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC	activity, and can be used in gene therapy and vaccine production. (I)
CC	proteins and polynucleotides may be used in the prevention, diagnosis and
CC	treatment of diseases associated with inappropriate (I) expression. For
CC	example, they may be used to treat disorders associated with decreased
CC	expression by rectifying mutations or deletions in a patient's genome
CC	that affect the activity of (I) by expressing inactive proteins or to
CC	supplement the patients own production of (I). Additionally, (I)
CC	polynucleotides may be used to produce the secreted (I), by inserting the
CC	nucleic acids into a host cell and culturing the cell to express the
CC	protein. (I) proteins and polynucleotides may be used to prevent,
CC	diagnose and treat immune/haematopoietic-related diseases, especially
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC	to AAK87694 represent human immune/haematopoietic antigen genomic
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC	represent sequences used in the exemplification of the present invention
XX	
SEQ	Sequence 1255 BP; 372 A; 274 G; 270 G; 339 T; 0 U; 0 Other;
Query Match	1.2%; Score 25; DB 4; Length 1255;
Best Local Similarity	100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches	0; Indels 0; Gaps 0
OY	2012 CTACTAAATAATATAATTAAGCTG 2036
Db	1087 CTACTAAATAATATAATTAAGCTG 1111
RESULT 83	
ABV25475/C	
ID	ABV25475 standard; cDNA; 1265 BP.
XX	
AC	ABV25475;
XX	
DT	16-SEP-2002 (first entry)
XX	
DE	Human prostate expression marker cDNA 25466.
XX	
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX	pharmacogenomic marker; gene; ss.
OS	Homo sapiens.
XX	
PN	WO200160860-A2.
XX	
PD	23-AUG-2001.
XX	
PF	20-FEB-2001; 2001WO-US005171.
XX	
PR	17-FEB-2000; 2000US-0183319P.
XX	
PR	16-MAR-2000; 2000US-0189862P.
XX	
PR	25-MAY-2000; 2000US-0207454P.
XX	
PR	09-JUN-2000; 2000US-0211314P.
XX	
PR	18-JUL-2000; 2000US-0218007P.
XX	
PR	13-DEC-2000; 2000US-0255281P.
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX	
PI	Schlegel R, Endege WO, Monahan JE;
XX	
DR	WPI; 2001-662795/76.
XX	
PT	Novel isolated nucleic acid molecule associated with cancerous state of
XX	prostate cells and correlating with presence of prostate cancer, useful
PT	for detecting presence of prostate cancer, stage of prostate cancer.
XX	
PS	Claim 1; Page 5059; 11750pp; English.
XX	
CC	The invention relates to an isolated nucleic acid molecule (I) comprising
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (i) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
 CC

Sequence 1265 BP; 371 A; 268 C; 254 G; 361 T; 0 U; 11 Other;

Query Match 1.2%; Score 25; DB 5; Length 1265;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTTACTAAATATATAAATTAGCTG 2036  
 |||||  
 369 CTTACTAAATATATAAATTAGCTG 345

RESULT 84  
 AAH18060  
 ID AAH18060 standard; cDNA; 1289 BP.  
 AC AAH18060;  
 DT 26-JUN-2001 (first entry)  
 XX  
 XX Human cDNA sequence SEQ ID NO:17895.  
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KM Homo sapiens.  
 OS

XX  
 XX EPI074617-A2.  
 PN  
 XX  
 XX 07-FEB-2001.  
 PD  
 XX  
 XX 28-JUL-2000; 2000EP-00116126.  
 PF  
 XX  
 XX 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 XX WPI; 2001-318749/34.  
 DR  
 XX  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.

PS Claim 8; SEQ ID NO 17895; 2537pp + Sequence Listing; English.  
 XX  
 XX The present invention describes primer sets for synthesizing 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 CC

Sequence 1289 BP; 430 A; 234 C; 265 G; 360 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 4; Length 1289;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTTACTAAATATATAAATTAGCTG 2036  
 |||||  
 1127 CTTACTAAATATATAAATTAGCTG 1151

RESULT 85  
 ABZ71641/C  
 ID ABZ71641 standard; cDNA; 1382 BP.  
 XX  
 XX ABZ71641;  
 AC  
 XX  
 XX 04-APR-2003 (first entry)  
 DT  
 XX  
 XX Breast specific nucleic acid # SEQ ID 101.  
 DE  
 XX  
 XX Human; breast specific nucleic acid; breast cancer; gene therapy;  
 KM cytostatic; ss.  
 KW  
 XX  
 XX Homo sapiens.  
 OS

XX  
 XX WO200288375-A2.  
 PN  
 XX  
 XX 07-NOV-2002.  
 PD  
 XX  
 XX 29-OCT-2001; 2001WO-US051318.  
 PF  
 XX  
 XX 27-OCT-2000; 2000US-0243805P.  
 PR  
 XX  
 XX (DIAD-) DIADEXUS INC.  
 PA  
 XX  
 XX Sun Y, Recipon H, Salceda S, Liu C, Turner LR;  
 PI  
 XX  
 XX WPI; 2003-156692/15.  
 DR  
 XX  
 XX  
 PT New breast-specific nucleic acids and polypeptides, useful for  
 PT identifying, diagnosing, monitoring, staging, imaging, and treating  
 PT breast cancer and non-cancerous disease states in breast tissues.

PS 1 b; Page 213; 269pp; English.  
 XX  
 XX The invention relates to breast-specific nucleic acid and polypeptide  
 CC sequences. The activity of sequences of the invention may be described as  
 CC cytostatic. The breast-specific nucleic acids, polypeptides and  
 CC compositions comprising them are useful for identifying, diagnosing,  
 CC monitoring, staging, imaging, and treating breast cancer and non-  
 CC cancerous disease states in breast tissue. They are also useful for  
 CC identifying breast tissue, for monitoring, identifying or designing  
 CC agonists and antagonists of the polypeptides, in gene therapy, in  
 CC producing transgenic animals and cells, for producing engineered breast  
 CC tissue for treatment and research, and as elements in an array or  
 CC computer program for pattern recognition of breast disorders. The nucleic  
 CC acids may be used as hybridisation probes to detect, characterise and  
 CC quantify hybridising nucleic acids in, and isolate hybridising nucleic  
 CC acids from, both genomic and transcript-derived nucleic acid samples. The  
 CC sequences given in records ABZ71541-ABZ71693 represent breast-specific

CC nucleic acid sequences of the invention  
XX Sequence 1382 BP; 432 A; 269 C; 323 G; 358 T; 0 U; 0 Other;  
SQ

Query Match 1.2%; Score 25; DB 8; Length 1382;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAAATATATAAATAGCTG 2036  
Db 705 CTAATAAAATATATAAATAGCTG 681

RESULT 86  
ADP85883/c  
ID ADP85883 standard; cDNA; 1382 BP.  
XX  
AC ADP85883;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Human breast specific nucleic acid (BSNA) cDNA, SEQ ID NO:101.  
XX  
KW Human; breast specific nucleic acid; BSNA; breast cancer; metastasis;  
KW diagnosis; monitoring; staging; imaging; immunotherapy; vaccine;  
KW cytostatic; gene therapy; transgenic animal; tissue engineering; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003099989-A2.  
XX  
PD 04-DEC-2003.  
XX  
PF 22-MAY-2002; 2002WO-US016307.  
XX  
PR 22-MAY-2002; 2002WO-US016307.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Sun Y, Recljon H, Salceda S, Liu C, Turner LR;  
XX  
DR WPI; 2004-042804/04.  
XX  
PT New breast specific nucleic acid molecules and proteins, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating breast  
PT cancer and non-cancerous disease states in breast tissue.  
XX  
PS Claim 1; SEQ ID NO 101; 269bp; English.  
XX  
XX The invention relates to 153 breast specific nucleic acids (BSNA;  
CC ADP85783-ADP85935) and to breast specific proteins (BSP; ADP85936-  
CC ADP86004). The invention also encompasses sequences at least 60%  
CC identical to the BSNA and BSPs of the invention; vectors and host cells  
CC comprising a BSNA; the recombinant production of BSPs using a host cell  
CC comprising a BSNA; an antibody specific for a BSP; methods of detecting  
CC an BSNA or BSP; a method for diagnosing or monitoring the presence and  
CC metastases of breast cancer in a patient; a kit for detecting the risk or  
CC presence of cancer in a patient; a method for the treatment of breast  
CC cancer via the administration of a BSNA or BSP to raise an immune  
CC response against the breast cancer cell expressing the BSNA or BSP; and a  
CC vaccine composition comprising a BSNA or BSP. The BSNA, BSPs, methods  
CC and compositions of the present invention are useful for diagnosing,  
CC monitoring, staging, imaging, and treating breast cancer and breast  
CC cancer metastases, and also non-cancerous disease states in breast  
CC tissue. The BSNA and BSPs are also useful for producing transgenic  
CC animals and cells, for producing engineered breast tissue for treatment  
CC and research, or for identifying breast tissue for breast cell  
CC differentiation and development. The present sequence represents a  
CC specifically claimed breast specific nucleic acid (BSNA). Note: 113 BSPs  
CC are claimed in the invention; however, only 69 of these (ADP85936-  
CC ADP86004) are given in the specification.  
XX  
SQ Sequence 1382 BP; 432 A; 269 C; 323 G; 358 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 12; Length 1382;  
Best Local Similarity 100.0%; Pred. No. 0.38; 0; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 2012 CTAATAAAATATATAAATAGCTG 2036  
Db 705 CTAATAAAATATATAAATAGCTG 681

RESULT 87  
AAD12601/c  
ID AAD12601 standard; cDNA; 1387 BP.  
XX  
AC AAD12601;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Human protein having hydrophobic domain encoding cDNA clone HP10794.  
XX  
KW Human; hydrophobic domain; gene therapy; nutritional supplement;  
KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;  
KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;  
KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;  
KW Huntington's disease; Alzheimer's disease; chemokine; chemokine;  
KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;  
KW contraceptive; antifertility; antiinflammatory; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FH CDS 147..488  
FT /\*tag= a  
FT /product= "Human protein having hydrophobic domain"  
FT /note= "CDS is specifically is claimed in claim 3"  
XX  
PN WO200149728-A2.  
XX  
PD 12-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-JP009359.  
XX  
PR 06-JAN-2000; 2000JP-00000585.  
XX  
PR 06-JAN-2000; 2000JP-00000588.  
XX  
PR 11-JAN-2000; 2000JP-00002299.  
XX  
PR 03-FEB-2000; 2000JP-00026862.  
XX  
PR 03-MAR-2000; 2000JP-00058367.  
XX  
XX (PROT-) PROTEGENE INC.  
XX  
PA (SAGA ) SAGAMI CHEM RES CENT.  
XX  
PI Kato S, Kimura T,  
XX  
DR WPI; 2001-418355/44.  
XX  
DR P-PSDB; AAE06606.  
XX  
PT Human proteins with hydrophobic domains and the nucleic acids encoding  
PT them, useful for preventing diagnosing and treating e.g. cancer,  
PT Alzheimer's and inflammation.  
XX  
XX  
XX Claim 4; Page 475-477; 563bp; English.  
XX  
XX The present sequence is human protein with hydrophobic domain encoding  
CC cDNA clone HP10794. The polynucleotide and polypeptide of the invention  
CC may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. The polynucleotides  
CC may be used to produce the polypeptide, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the protein. The  
CC polynucleotides and its complementary sequences may also be used as DNA  
CC probes in diagnostic assays and also used in gene therapy. The  
CC polypeptides may also be used as antigens in the production of antibodies  
CC and in assays to identify modulators of polypeptide expression and  
CC activity. The polypeptides and nucleic acids may be used as nutritional



[illegible]

Pt	Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC,
Pt	Bednarik DP, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM,
Pt	Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;
Pt	Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H,
Pt	WPI; 1998-609887/51.
Dn	P-PsDB; AAM75151.
Pt	New isolated human genes and the secreted polypeptides they encode -
Pt	useful for diagnosis and treatment of e.g. cancers, neurological
Pt	disorders, immune diseases, inflammation or blood disorders.
Pt	Claim 1; Page 248-249; 447pp; English.
Pt	This sequence represents a nucleic acid molecule which encodes a secreted
Pt	human protein. The gene number, and the clone it is derived from, are
Pt	detailed in the descriptor line. The gene can be used to generate fusion
Pt	proteins by linking to the gene to a human immunoglobulin Fc portion
Pt	(e.g. AAV34145) for increasing the stability of the fused protein as
Pt	compared to the human protein only. The invention relates to 70 novel
Pt	genes and their fragments (nucleic acid sequences: AAV34154-V34276; amino
Pt	acid sequences AAW75057-W75179) which are useful for preventing, treating
Pt	or ameliorating medical conditions e.g. by protein or gene therapy. Also,
Pt	pathological conditions can be diagnosed by determining the amount of the
Pt	new polypeptides in a sample or by determining the presence of mutations
Pt	in the new polynucleotides. Specific uses are described for each of the
Pt	70 polynucleotides, based on which tissues they are most highly expressed
Pt	in (see AAV34154 for described uses). (Updated on 25-MAR-2003 to correct
Pt	Pf field.) (Updated on 25-MAR-2003 to correct Pf field.)
Sq	Sequence 1822 BP; 528 A; 379 C; 349 G; 564 T; 0 U; 2 Other;
Qy	Query Match: 1.2%; Score 25; DB 2; Length 1822;
Qy	Best Local Similarity: 100.0%; Pred. No. 0.38;
Mt	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Oy	2012 CTACTAAAAATATATAAAATTAGCTG 2036       180 CTACTAAAATATATAAAATTAGCTG 156
Id	ACD08119/C
Ac	ACD08119 standard; cDNA; 1822 BP.
Ac	ACD08119;
Dt	12-AUG-2003 (First entry)
Dt	cDNA encoding novel human secreted protein #95.
Xx	Human; immunoglobulin G; IgG; fragment of crystallisation; Fc; Kw Immune system disorder; haematopoietic cell disorder; Kw immunologic deficiency disorder; ataxia telangiectasia; HIV infection; Kw Wiskott-Aldrich disorder; thrombocytopenia; haemoglobinuria; Kw blood coagulation disorder; blood platelet disorder; autoimmune disorder; Kw Addison's disease; haemolytic anaemia; rheumatoid arthritis; dermatitis; Kw glomerulonephritis; Grave's disease; allergic reaction; Kw great-vessels-host disease; hyperproliferative disorder; neoplasia; Kw infectious disease; nervous system disease; spinal cord disorder; Kw head trauma; stroke; tissue regeneration; congenital defect; trauma; Kw wound; burn; incision; ulcer; age disease; osteoporosis; Kw periodontal disease; liver failure; cataplexy; anabolism; metabolism; Kw food additive; preservative; secreted protein; gene; ss. Xx Homo sapiens. Xx US2003027132-A1. Xx 06-FEB-2003. Xx PD 04-SEP-1998; 98US-00148545. Xx PF



XX  
PR 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040161P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040526P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043368P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043659P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 22-AUG-1997; 97US-0056630P.  
PR 22-AUG-1997; 97US-0056631P.  
PR 22-AUG-1997; 97US-0056632P.  
PR 22-AUG-1997; 97US-0056633P.  
PR 22-AUG-1997; 97US-0056636P.  
PR 22-AUG-1997; 97US-0056637P.  
PR 22-AUG-1997; 97US-0056662P.  
PR 22-AUG-1997; 97US-0056664P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056862P.  
PR 22-AUG-1997; 97US-0056864P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056874P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056888P.  
PR 22-AUG-1997; 97US-0056889P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057761P.  
PR 06-MAR-1998; 98WO-US004482.  
  
XX  
PA (RUBE/) RUBEN S M.  
PA (ROSE/) ROSEN C A.  
PA (FISC/) FISCHER C L.  
PA (SOPP/) SOPPET D R.  
PA (CART/) CARTER K C.  
PA (BEDN/) BEDNARIK D R.  
PA (ENDR/) ENDRESS G A.  
PA (YUGG/) YU G.  
PA (NIJ/) NI J.  
PA (FENG/) FENG P.  
PA (YOUN/) YOUNG P E.  
PA (GREE/) GREENE J M.  
PA (FERR/) FERRIE A M.  
PA (DUAN/) DUAN R.  
PA (HUJ/) HU J.  
PA (FLOF/) FLORENCE K A.  
PA (OLSE/) OLSEN H S.  
PA (EBNE/) EBNER R.  
PA (BREW/) BREWER L A.  
PA (SHIY/) SHI Y.  
  
XX  
PI Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC;  
PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM;  
PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;  
PI Shi Y;  
  
XX  
DR WPI: 2003-466138/44.  
DR P-PSDB: ABO02027.  
  
XX  
PT Novel isolated human secreted HODAZ50 polypeptide useful for diagnosing  
PT or treating deficiencies or disorders of the immune system, autoimmune  
PT disorders, hyperproliferative disorders, and infectious diseases.  
  
XX  
PS Claim 4; Page 145-146; 243pp; English.  
  
XX  
CC The invention describes an isolated human secreted HODAZ50 polypeptide  
CC (I) comprising a sequence at least 95% identical to a sequence selected  
CC from polypeptide fragment of any one of the 123 polypeptide sequences  
CC (PS) fully defined in the specification and having biological activity,  
CC polypeptide domain or epitope of PS, secreted form of PS, full-length  
CC protein of PS, or variant, allelic variant or species homologue of PS.  
CC (I) or a polynucleotide (II) encoding (I) is useful for preventing,  
CC treating, or ameliorating a medical condition in a mammalian subject. (I)  
CC or (II) is also useful for diagnosing a pathological condition or a  
CC susceptibility to a pathological condition in a subject. (I) is useful  
CC for identifying a binding partner which involves contacting the  
CC polypeptide with the binding partner and determining whether the binding  
CC partner affects the activity of the polypeptide. (I) or (II) is useful  
CC for diagnosing or treating deficiencies or disorders of the immune  
CC system, deficiencies or disorders of haematopoietic cells, to treat

CC immunologic deficiency disorders, ataxia telangiectasia, HIV infection,  
CC Wiskott-Aldrich disorders, thrombocytopenia or haemoglobinuria, blood  
CC coagulation disorders, blood platelet disorders, autoimmune disorders  
CC (e.g., Addison's disease, hemolytic anaemia, rheumatoid arthritis,  
CC dermatitis, glomerulonephritis, Grave's disease), allergic reactions,  
CC graft-versus-host disease, hyperproliferative disorders (e.g., neoplasms  
CC located in the abdomen, bone, breast, digestive system, liver, pancreas,  
CC peritoneum, endocrine glands), infectious diseases (e.g., viral,  
CC bacterial, fungal or parasitic infection) central and peripheral nervous  
CC system diseases (e.g., spinal cord disorders, head trauma or stroke), to  
CC differentiate, proliferate and attract cells leading to the regeneration  
CC of tissues to repair, replace or protect tissue damaged by congenital  
CC defects, trauma (wounds, burns, incisions, or ulcers) age disease (e.g.,  
CC osteoporosis, periodontal disease, liver failure) or surgery. (I) or (IV)  
CC is useful to mediate mammalian characteristics, to modulate mammalian  
CC metabolism affecting catabolism, anabolism, processing, utilisation, and  
CC storage of energy, to change a mammal's mental state or physical state,  
CC or as a food additive or preservative, such as to increase or decrease  
CC storage capabilities, fat content, lipid, protein, carbohydrate,  
CC vitamins, minerals, cofactors or other nutritional components. This  
CC sequence encodes a novel human secreted protein

Seq Sequence 1822 BP, 528 A, 379 C, 349 G, 564 T, 0 U, 2 Other;

Query Match 1.2%; Score 25; DB 8; Length 1822;

Best Local Similarity 100.0%; Pred. No. 0.38;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTTAAATATTAATAATTAGCTG 2036

DB 180 CTACTTAAATATTAATAATTAGCTG 156

RESULT 91

AAS26824 AAS26824 standard; cDNA; 1859 BP.

AC AAS26824;

DT 07-NOV-2001 (first entry)

DE Human cDNA encoding a novel secreted protein, SEQ ID 16.

XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
XX cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective;  
XX antibacterial; virucide; fungicide; ophtalmological; vulnereary;  
XX secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
XX cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
XX cerebral ischaemia; angiogenesis; nervous system disorder;  
XX Alzheimer's disease; infection; ocular disorder; corneal infection;  
XX wound healing; epithelial cell proliferation; skin ageing; food additive;  
XX preservative; antiproliferative.

OS Homo sapiens.

PN WO200155441-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001320.

XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217490P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227099P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232403P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-476222/51.  
DR P-PSDB; AAU16919.  
XX  
XX  
PT Novel polypeptides and polynucleotides useful as diagnostic reagents to  
PT diagnose diseases or disorders associated with aberrant expression or  
PT activity of polypeptides, for treating blood clotting disorder,  
PT hemophilia.  
XX  
XX  
XX Claim 1; SEQ ID NO 16; 601pp; English.  
XX  
XX  
CC The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic

CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. Rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention. Note: The

Query Match 1.2% Score 25; DB 4; Length 1859;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAAATTACCTG 2036  
Db 1673 CTACTAAAAATATATAAAATTACCTG 1697

RESULT 92  
AA528995  
ID AA528995 standard; DNA; 1861 BP.  
XX  
AC AA528995;  
DT 21-NOV-2001 (first entry)  
XX  
DE Genomic sequence #1 encoding novel human uterine motility polypeptide.  
XX  
DE Human; uterine motility-association disorder; uterus; pregnancy; labour;  
XX menstrual cycle; gene therapy; ds.  
XX  
XX Homo sapiens.  
OS  
PN WO200155201-A1.  
XX  
PD 02-AUG-2001.  
XX  
XX  
PF 17-JAN-2001; 2001WO-US001317.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239835P.  
PR 13-OCT-2000; 2000US-0239837P.  
PR 20-OCT-2000; 2000US-0240963P.  
PR 20-OCT-2000; 2000US-0241211P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 05-DEC-2000; 2000US-0251990P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251859P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488777/53.

PT New nucleic acid molecules encoding 49 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.

XX Disclosure; SEQ ID NO 129; 524bp; English.

XX The present invention relates to the isolation of novel human uterine  
XX motility-association polypeptides (AAU18094-AAU18152), and cDNA and  
XX genomic sequences encoding for these polypeptides. The sequences of the  
XX invention are useful in the diagnosis, treatment, prevention and/or  
XX prognosis of diseases associated with uterine motility such as pregnancy  
XX and labour, and menstrual disorders. The polynucleotide sequences of the  
XX invention are also useful in gene therapy. AAS28995-AAS29020 represent  
XX genomic sequences encoding for novel human uterine motility-association  
XX polypeptides. Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic format  
XX directly from WIPO at fcp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1861 BP; 411 A; 479 C; 591 G; 380 T; 0 U; 0 Other;

XX Query Match 1.2%; Score 25; DB 4; Length 1861;

XX Best Local Similarity 100.0%; Pred. No. 0.38; Mismatches 0; Indels 0; Gaps 0;

XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATTTAAATTTAGCTG 2036

DB 1674 CTTACTAAAAATTTAAATTTAGCTG 1698

```
RESULT 93
AAS31660
ID AAS31660 standard; DNA; 1861 BP.
XX
AC AAS31660;
XX
DT 04-DEC-2001 (first entry)
XX
DE Genomic sequence #6 encoding novel human calcium-binding protein.
XX
KW Human; calcium-binding protein; calcium flux; neurological disease;
KW immune dysfunction; digestive disorder; neoplastic disease;
KW blood disorder; infectious disease; gene therapy; immunosuppressive;
KW antidiarrhetic; cytoskeletal; vasotropic; antibacterial; nootropic;
KW virucide; ds.
XX
OS Homo sapiens.
XX
PN WO200155304-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001302.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
```

PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-465568/50.  
XX  
XX  
PT Isolated nucleic acid molecule encoding a calcium-binding protein is used  
PT in preventing, treating or ameliorating a medical condition.  
XX  
PS Disclosure; SEQ ID NO 172; 542bp; English.  
XX  
XX The present invention relates to the isolation of novel human calcium-  
CC binding proteins (AAU9892-AAU9969), and cDNA and genomic sequences  
CC encoding for these proteins. The sequences of the invention are useful in  
CC the diagnosis, prevention and/or prognosis of diseases associated with  
CC aberrant calcium flux. Such disorders include neurological diseases (e.g.  
CC amyotrophic lateral sclerosis, ALS), immune dysfunction (e.g. severe  
CC combined immunodeficiency, SCID), digestive disorders (e.g. irritable  
CC bowel syndrome, IBS), neoplastic disease (e.g. cancer), blood disorders  
CC (e.g. hemophilia), and/or infectious disease (e.g. acquired  
CC immunodeficiency syndrome, AIDS). The novel calcium-binding proteins are  
CC also useful as screening tools to identify antagonists and/or agonists  
CC that may enhance or inhibit activities mediated by calcium-binding  
CC proteins. The polynucleotides of the invention are also useful in gene  
CC therapy. AAS1655-AAS1680 represent genomic sequences encoding for the  
CC novel human calcium-binding proteins. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1861 BP; 411 A; 479 C; 591 G; 380 T; 0 U; 0 Other;  
Query Match 1.2%; Score 25; DB 4; Length 1861;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTAATAAAATATAAAATTAGCTG 2036  
DB 1674 CTAATAAAATATAAAATTAGCTG 1698  
RESULT 94  
ABT07830  
ID ABT07830 standard; DNA; 1861 BP.  
XX  
AC ABT07830;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Novel human nucleic acid SEQ ID No 129.  
XX  
XX Immunostimulant; antineumatic; antiarthritic; neuroprotective;  
XX anti-allergic; antidiabetic; antiaslathmic; antiinflammatory; nootropic;  
XX immunosuppressive; anticoagulant; thrombolytic; antithrombotic;  
XX cytosolic; nephrotropic; antiparkinsonian; gynecological; vituicide;  
XX antibacterial; antirhythmic; fungicide; HCFAT05; HMAAB95; HTNEM01;  
XX immunodeficiency; autoimmune disorder; allergic reaction; cardiovascular;  
XX inflammatory condition; graft-versus-host disease; reproductive system;

KM blood-related disorder; hyperproliferative; endocrine; neurological;  
KM respiratory; renal; infectious disease; gastrointestinal; gene therapy;  
KM neuronal growth; neuronal disorder; neuro-degenerative condition;  
KM keratinocyte growth; human; ds.  
OS Homo sapiens.  
XX  
XX US2002086330-A1.  
XX  
PD 04-JUL-2002.  
XX  
XX  
XX 17-JAN-2001; 2001US-00764893.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225470P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 22-AUG-2000; 2000US-0226868P.  
XX 30-SEP-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 08-SEP-2000; 2000US-0231413P.  
XX 21-SEP-2000; 2000US-0234223P.  
XX 21-SEP-2000; 2000US-0234274P.  
XX 25-SEP-2000; 2000US-0234997P.  
XX 27-SEP-2000; 2000US-0235834P.  
XX 29-SEP-2000; 2000US-0236327P.  
XX 29-SEP-2000; 2000US-0236367P.  
XX 29-SEP-2000; 2000US-0236368P.  
XX 29-SEP-2000; 2000US-0236369P.  
XX 29-SEP-2000; 2000US-0236370P.  
XX 02-OCT-2000; 2000US-0236802P.  
XX 02-OCT-2000; 2000US-0237037P.  
XX 02-OCT-2000; 2000US-0237038P.  
XX 02-OCT-2000; 2000US-0237039P.  
XX 02-OCT-2000; 2000US-0237040P.  
XX 13-OCT-2000; 2000US-0239335P.  
XX 20-OCT-2000; 2000US-0240960P.  
XX 20-OCT-2000; 2000US-0241785P.  
XX 20-OCT-2000; 2000US-0241809P.  
XX 01-NOV-2000; 2000US-0244617P.  
XX 17-NOV-2000; 2000US-0249299P.  
XX 08-DEC-2000; 2000US-0251856P.  
XX 08-DEC-2000; 2000US-0251868P.  
XX 08-DEC-2000; 2000US-0251869P.  
XX  
XX (ROSE/) ROSEN C A.  
XX (RUBE/) RUBEN S M.  
XX (BARA/) BARASH S C.  
XX  
XX Rosen CA, Ruben SM, Barash SC,  
XX  
XX WPI; 2002-665432/71.  
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and

PT treatment of immune, hyperproliferative, renal, respiratory, cardiovascular, reproductive, endocrine, gastrointestinal and neurological disorders.

PS Disclosure; Page 295-296; 335pp; English.

CC The invention relates to an isolated polypeptide comprising a sequence at  
CC least 90% identical to a full length protein sequence selected from 55  
CC sequences given in the specification such as a sequence of 163, 74 or 140  
CC amino acids fully defined in the specification, or the encoding sequence  
CC contained in 49 cDNA clones given in specification, e.g. HCRAT05, HMAA95  
CC or HTMEM01. The protein and its encoding nucleic acid are useful for  
CC diagnosing a pathological condition or susceptibility to a pathological  
CC condition in a subject and for preventing, treating or ameliorating a  
CC medical condition. The protein, its encoding nucleic acid and an isolated  
CC antibody that can bind to the protein are useful in treating, preventing,  
CC diagnosing and/or prognosing immunodeficiencies, autoimmune disorders,  
CC allergic reactions and conditions, inflammatory conditions, graft-versus-  
CC host disease, blood-related disorders, hyperproliferative disorders,  
CC renal disorders, cardiovascular disorders, respiratory disorders,  
CC neurological disorders, endocrine disorders, reproductive system  
CC disorders, infectious diseases, and gastrointestinal disorders. The  
CC protein of the invention is useful to stimulate neuronal growth and to  
CC treat, prevent, and/or diagnose neuronal damage which occurs in certain  
CC neuronal disorders or neuro-degenerative conditions, for stimulating  
CC keratinocyte growth, to prevent hair loss, to modulate mammalian  
CC characteristics such as body height, weight, hair color, and to increase  
CC or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrate, vitamins, minerals, cofactors or other nutritional  
CC components. The nucleic acid of the invention can be used in gene  
CC therapy. This polynucleotide sequence represents one of the novel nucleic  
CC acids of the invention

SQ Sequence 1861 BP; 411 A; 479 C; 591 G; 380 T; 0 U; 0 Other;

Query Match	1.2%	Score 25;	DB 6;	Length 1861;
Best Local Similarity	100.0%;	Pred. No. 0.38;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

<b>QY</b>	2012	CTACTAAAAATATATAAAATTAGCTG	2036
<b>Db</b>	1674	CTACTAAAAATATATAAAATTAGCTG	1698

RESULT 95  
ABA06374  
ID ABA06374 standard; cDNA; 1869 BP.  
...

PR	18-MAY-2000	2000US-019813.3
PR	9-MAY-2000	2000US-020551.7
PR	07-JUN-2000	2000US-020946.7
PR	28-JUN-2000	2000US-021488.6
PR	30-JUN-2000	2000US-021513.5
PR	07-JUL-2000	2000US-021664.7
PR	07-JUL-2000	2000US-021688.0
PR	11-JUL-2000	2000US-021748.7
PR	11-JUL-2000	2000US-021748.6
PR	14-JUL-2000	2000US-021829.0
PR	26-JUL-2000	2000US-022096.3
PR	14-AUG-2000	2000US-022451.8
PR	14-AUG-2000	2000US-022451.9
PR	14-AUG-2000	2000US-022521.3
PR	14-AUG-2000	2000US-022521.4
PR	14-AUG-2000	2000US-022526.6
PR	14-AUG-2000	2000US-022526.7
PR	14-AUG-2000	2000US-022526.8
PR	14-AUG-2000	2000US-022527.0
PR	14-AUG-2000	2000US-022544.7
PR	14-AUG-2000	2000US-022575.7
PR	14-AUG-2000	2000US-022575.8
PR	14-AUG-2000	2000US-022575.9
PR	18-AUG-2000	2000US-022627.9
PR	22-AUG-2000	2000US-022668.1
PR	22-AUG-2000	2000US-022686.8
PR	23-AUG-2000	2000US-022718.2
PR	23-AUG-2000	2000US-022700.9
PR	30-AUG-2000	2000US-022897.4
PR	01-SEP-2000	2000US-022928.7
PR	01-SEP-2000	2000US-022934.3
PR	01-SEP-2000	2000US-022934.4
PR	01-SEP-2000	2000US-022934.5
PR	05-SEP-2000	2000US-022950.9
PR	05-SEP-2000	2000US-022951.3
PR	06-SEP-2000	2000US-023043.7
PR	06-SEP-2000	2000US-023043.8
PR	08-SEP-2000	2000US-023124.2
PR	08-SEP-2000	2000US-023124.3
PR	08-SEP-2000	2000US-023124.4
PR	08-SEP-2000	2000US-023141.3
PR	08-SEP-2000	2000US-023141.4
PR	08-SEP-2000	2000US-023208.0
PR	08-SEP-2000	2000US-023208.1
PR	12-SEP-2000	2000US-023196.8
PR	14-SEP-2000	2000US-023238.7
PR	14-SEP-2000	2000US-023238.8
PR	14-SEP-2000	2000US-023239.9
PR	14-SEP-2000	2000US-023240.1
PR	14-SEP-2000	2000US-023240.2
PR	14-SEP-2000	2000US-023306.3
PR	14-SEP-2000	2000US-023306.4
PR	14-SEP-2000	2000US-023306.5
PR	21-SEP-2000	2000US-023422.3
PR	21-SEP-2000	2000US-023422.4
PR	25-SEP-2000	2000US-023498.7
PR	25-SEP-2000	2000US-023498.8
PR	26-SEP-2000	2000US-023546.4
PR	27-SEP-2000	2000US-023583.6
PR	27-SEP-2000	2000US-023583.6
PR	29-SEP-2000	2000US-023633.7
PR	29-SEP-2000	2000US-023636.7
PR	29-SEP-2000	2000US-023636.8
PR	29-SEP-2000	2000US-023656.9
PR	29-SEP-2000	2000US-023637.0
PR	02-OCT-2000	2000US-023703.7
PR	02-OCT-2000	2000US-023703.8
PR	02-OCT-2000	2000US-023703.9
PR	02-OCT-2000	2000US-023704.0
PR	13-OCT-2000	2000US-023993.5
PR	13-OCT-2000	2000US-023993.7

PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250316P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
DR WPI, 2001-476161/51.  
XX P-PSDB; ABB10152.  
XX  
PT Isolated nucleic acid molecule encoding an inflammation-associated  
PT polypeptide is used in preventing, treating or ameliorating a medical  
PT condition.  
XX  
XX  
PS Claim 1; SEQ ID NO 40; 859pp + Sequence Listing; English.  
XX  
CC The present invention provides human cDNAs, proteins and related genomic  
CC DNAs. These can be used in the treatment of neural, immune system,

CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal and proliferative disorders and inflammation. The present sequence  
CC is a cDNA of the invention  
XX  
SQ Sequence 1869 BP; 408 A; 480 C; 595 G; 379 T; 0 U; 7 Other;  
Query Match 1.2%; Score 25; DB 4; Length 1869;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTTAAATATTAATAATTAGCTG 2036  
DB 1673 CTACTTAAATATTAATAATTAGCTG 1697  
RESULT 96  
ABV83711  
ID ABV83711 standard; cDNA, 1869 BP.  
XX  
AC ABV83711;  
XX  
DT 09-DEC-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 40.  
XX  
KW Human; nocotropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
KW gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
PN US2002090672-A1.  
XX  
PD 11-JUL-2002.  
XX  
XX 17-JAN-2001; 2001US-00764853.  
PF  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225577P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.



PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 01-NOV-2000; 2000US-0241809P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

DR WPI; 2002-68127/73.  
P-PSDB; ABP66739.

PT Novel polypeptide useful for diagnosis, prognosis, prevention, and  
PT treatment of immune, hyperproliferative, renal, respiratory,  
PT cardiovascular, reproductive, endocrine, gastrointestinal and  
PT neurological disorders.

PS Claim 1; SEQ ID NO 40; 369pp + Sequence Listing; English.

CC The invention relates to novel genes (ABV83682-ABV84101) and proteins  
CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, diabetes mellitus, Crohn's disease,  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1869 BP; 408 A; 480 C; 595 G; 379 T; 0 U; 7 Other;

Query Match 1.2%; Score 25; DB 6; Length 1869;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTAAGTAAATTAATTAATTAAGCTG 2036  
DB 1673 CTAAGTAAATTAATTAATTAAGCTG 1697

RESULT 97

AAH15709  
ID AAH15709 standard; cDNA; 1880 BP.

XX AAH15709;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:14088.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300252.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PS Claim 8; SEQ ID NO 14088; 2537pp + Sequence Listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

XX SQ Sequence 1880 BP; 532 A; 392 C; 438 G; 518 T; 0 U; 0 Other;

Query Match 1.2%; Score 25; DB 4; Length 1880;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTAAGTAAATTAATTAATTAAGCTG 2036  
DB 1721 CTAAGTAAATTAATTAATTAAGCTG 1745

RESULT 98

AAH16703  
ID AAH16703 standard; cDNA; 1947 BP.

XX AAH16703;

DT 26-JUN-2001 (first entry)

```

XX Human cDNA sequence SEQ ID NO:15869.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX
XX PD 07-FEB-2001.
XX
XX PF 28-JUL-2000; 2000EP-00116126.
XX
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX
XX PA (HELI-) HELIX RES INST.
XX
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI; 2001-318749/34.
XX
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX
XX PS Claim 8; SEQ ID NO 15869; 2537pp + Sequence Listing; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification, where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX
XX SQ Sequence 1947 BP; 500 A; 446 C; 491 G; 510 T; 0 U; 0 Other;

```

```

Query Match 1.2%; Score 25; DB 4; Length 1947;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 1781 CTACTAAAAATATATAAATTAGCTG 1805

```

```

RESULT 99
ACA64683/c
ID ACA64683 standard; DNA; 1993 BP.
XX
XX AC ACA64683;
XX

```

```

DT 24-JUN-2003 (first entry)
XX
XX DE DNA encoding novel protein 158P3D2 var 5a.
XX
XX KW 158P3D2; cytostatic; gene therapy; vaccine; cancer; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200283928-A2.
XX
XX PD 24-OCT-2002.
XX
XX PF 25-MAR-2002; 2002WO-US009403.
XX
XX PR 10-APR-2001; 2001US-0283112P.
XX PR 25-APR-2001; 2001US-0286630P.
XX
XX PA (AGEN-) AGENSYS INC.
XX
XX PI Jakobovits A, Faris M, Morrison K, Morrison RK, Hubert RS;
XX PI Afar DEH, Ge W, Raitano AB, Challita-Bid PM;
XX DR WPI; 2003-167092/16.
XX DR P-PSDB; ABU78989.
XX
XX PT New composition comprising a substance that modulates the status of
XX PT 158P3D2 or a molecule that is modulated by 158P3D2, useful for treating
XX PT cancer.
XX
XX PS Example 2; Fig 2F; 354pp; English.
XX
XX CC The invention describes a new composition comprising a substance that
XX CC modulates the status of 158P3D2 or a molecule that is modulated by
XX CC 158P3D2, where the status of a cell that expresses 158P3D2 is modulated.
XX CC The composition is useful for treating cancer. This sequence encodes a
XX CC variant of the novel protein 158P3D2
XX
XX SQ Sequence 1993 BP; 422 A; 568 C; 538 G; 465 T; 0 U; 0 Other;

```

```

Query Match 1.2%; Score 25; DB 10; Length 1993;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 323 CTACTAAAAATATATAAATTAGCTG 299

```

```

RESULT 100
ACA64684/c
ID ACA64684 standard; DNA; 1993 BP.
XX
XX AC ACA64684;
XX
XX DT 24-JUN-2003 (first entry)
XX
XX DE DNA encoding novel protein 158P3D2 var 5b.
XX
XX KW 158P3D2; cytostatic; gene therapy; vaccine; cancer; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200283928-A2.
XX
XX PD 24-OCT-2002.
XX
XX PF 25-MAR-2002; 2002WO-US009403.
XX
XX PR 10-APR-2001; 2001US-0283112P.
XX PR 25-APR-2001; 2001US-0286630P.
XX
XX PA (AGEN-) AGENSYS INC.
XX

```

PI Jakobovits A, Faris M, Morrison K, Morrison RK, Hubert RS;  
 PI Afar DEH, Ge W, Raitano AB, Challita-Eid PM;  
 XX  
 DR WPI; 2003-167092/16.  
 DR P-PSDB; ABU78990.  
 XX  
 PT New composition comprising a substance that modulates the status of  
 PT 158P3D2 or a molecule that is modulated by 158P3D2, useful for treating  
 PT cancer.  
 XX  
 XX Example 2; Fig 2G; 354pp; English.  
 PS  
 CC The invention describes a new composition comprising a substance that  
 CC modulates the status of 158P3D2 or a molecule that is modulated by  
 CC 158P3D2, where the status of a cell that expresses 158P3D2 is modulated.  
 CC The composition is useful for treating cancer. This sequence encodes a  
 CC variant of the novel protein 158P3D2  
 CC  
 XX  
 SQ Sequence 1993 BP; 422 A; 568 C; 538 G; 465 T; 0 U; 0 Other;  
  
 Query Match 1.2%; Score 25; DB 10; Length 1993;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 2012 CTACTAAATAATATAAATTAGCTG 2036  
 |||||||  
 DB 323 CTACTAAATAATATAAATTAGCTG 299

Search completed: January 4, 2005, 18:39:24  
 Job time : 952 secs

***This Page Blank (uspio)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: January 4, 2005, 18:04:59 ; Search time 192 Seconds  
(without alignments)  
7537.324 Million cell updates/sec

Title: US-09-508-658B-1

Perfect score: 2036 1 agaccgggagagacggcgggg.....aaaataataaatactgctg 2036

Scoring table: OLIGO-NUC;  
Gapop 60.0 , Gapext 60.0

Searched: 824507 seqs, 355394441 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfillseq.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	1.2	172	4	US-09-513-999C-16265
2	25	1.2	239	4	US-09-513-999C-24183
3	25	1.2	413	4	US-09-470-276-64
4	25	1.2	486	4	US-09-621-976-14650
5	25	1.2	497	4	US-09-621-976-1876
6	25	1.2	530	4	US-09-621-976-1934
7	25	1.2	685	3	US-09-183-266A-16
8	25	1.2	1001	4	US-09-641-638-629
9	25	1.2	1001	4	US-09-641-638-630
10	25	1.2	1001	4	US-10-170-097-629
11	25	1.2	1001	4	US-10-170-097-630
12	25	1.2	1822	4	US-09-148-545-105
13	25	1.2	3694	3	US-09-232-200-46
14	25	1.2	3694	3	US-09-232-197-46
15	25	1.2	3694	3	US-09-232-201-46
16	25	1.2	3694	3	US-09-232-195-46
17	25	1.2	3694	3	US-09-232-195-46
18	25	1.2	3704	3	US-09-232-200-24
19	25	1.2	3704	3	US-09-232-197-24
20	25	1.2	3704	3	US-09-232-201-24
21	25	1.2	10380	3	US-09-077-354B-3
22	25	1.2	20674	4	US-09-641-638-651
23	25	1.2	20674	4	US-10-170-097-651
24	25	1.2	20966	4	US-09-776-976-7
25	25	1.2	20966	4	US-09-909-547-7
26	25	1.2	20966	4	US-09-569-852B-1
27	25	1.2	33769	4	US-09-544-398B-8

28	25	1.2	33769	4	US-09-543-771-8	Sequence 8, App1
29	25	1.2	40090	4	US-09-820-004-3	Sequence 3, App1
30	25	1.2	48763	4	US-09-916-204-3	Sequence 3, App1
31	25	1.2	48763	4	US-10-282-048-3	Sequence 3, App1
32	25	1.2	59065	3	US-09-813-817-3	Sequence 3, App1
33	25	1.2	59065	4	US-09-978-197-3	Sequence 3, App1
34	25	1.2	59065	4	US-10-135-696-3	Sequence 3, App1
35	25	1.2	62804	3	US-09-800-960-3	Sequence 3, App1
36	25	1.2	62804	4	US-10-096-960-3	Sequence 3, App1
37	25	1.2	75395	4	US-09-984-890-3	Sequence 3, App1
38	25	1.2	75395	4	US-10-274-194-3	Sequence 3, App1
39	25	1.2	114793	4	US-10-148-806-3	Sequence 3, App1
40	25	1.2	114793	4	US-10-148-806-3	Sequence 3, App1
41	25	1.2	137000	4	US-10-172-911-11	Sequence 11, App1
42	25	1.2	148567	4	US-09-801-876B-3	Sequence 3, App1
43	25	1.2	148567	4	US-10-254-869-3	Sequence 3, App1
44	25	1.2	174493	4	US-09-804-471A-3	Sequence 3, App1
45	25	1.2	174493	4	US-10-238-709-3	Sequence 3, App1
46	25	1.2	193303	4	US-09-497-855A-37	Sequence 37, App1
47	25	1.2	193303	4	US-09-497-855A-44	Sequence 44, App1
48	25	1.2	202001	4	US-09-734-674-3	Sequence 3, App1
49	25	1.2	246240	2	US-08-724-394A-20	Sequence 20, App1
50	25	1.2	246240	2	US-08-724-394A-21	Sequence 21, App1
51	25	1.2	246240	2	US-08-724-394A-22	Sequence 22, App1
52	25	1.2	392000	4	US-10-027-983-11	Sequence 11, App1
53	25	1.2	240	4	US-09-621-976-14785	Sequence 14785, A
54	25	1.2	444	4	US-09-621-976-14867	Sequence 14867, A
55	25	1.2	485	4	US-09-621-976-1665	Sequence 1665, Ap
56	25	1.2	491	4	US-09-621-976-1668	Sequence 1668, Ap
57	25	1.2	531	4	US-09-404-879A-24	Sequence 24, App1
58	25	1.2	531	4	US-09-338-933-24	Sequence 24, App1
59	25	1.2	531	4	US-09-215-681-24	Sequence 24, App1
60	25	1.2	531	4	US-09-216-003A-24	Sequence 24, App1
61	25	1.2	531	4	US-09-667-857-24	Sequence 24, App1
62	25	1.2	6678	3	US-08-816-617A-1	Sequence 1, App1
63	25	1.2	111822	3	US-09-754-250-3	Sequence 3, App1
64	25	1.1	163	4	US-09-513-999C-26863	Sequence 26863, A
65	25	1.1	175	4	US-09-513-999C-18352	Sequence 18352, A
66	25	1.1	288	1	US-08-410-804-22	Sequence 22, App1
67	25	1.1	288	1	US-08-259-514-22	Sequence 22, App1
68	25	1.1	316	2	US-08-858-311-22	Sequence 22, App1
69	25	1.1	316	4	US-09-513-999C-9318	Sequence 9318, Ap
70	25	1.1	382	4	US-09-621-976-14099	Sequence 14099, A
71	25	1.1	403	4	US-09-513-999C-31049	Sequence 31049, A
72	25	1.1	444	4	US-09-621-976-16675	Sequence 12675, A
73	25	1.1	694	1	US-08-708-107-1	Sequence 1, App1
74	25	1.1	803	4	US-09-495-050A-13	Sequence 13, App1
75	25	1.1	1001	4	US-09-671-314-150	Sequence 150, App
76	25	1.1	8133	3	US-09-659-791A-10	Sequence 10, App1
77	25	1.1	14636	4	US-09-173-914-6	Sequence 6, App1
78	25	1.1	15022	4	US-09-844-634-17	Sequence 17, App1
79	25	1.1	18400	4	US-09-901-151-3	Sequence 3, App1
80	25	1.1	19650	4	US-09-819-989-3	Sequence 3, App1
81	25	1.1	19650	4	US-10-273-992-3	Sequence 3, App1
82	25	1.1	35060	4	US-08-814-095-7	Sequence 7, App1
83	25	1.1	35060	4	US-09-740-041-3	Sequence 3, App1
84	25	1.1	66804	4	US-09-797-906-3	Sequence 3, App1
85	25	1.1	70000	4	US-09-851-896-3	Sequence 7, App1
86	25	1.1	72604	3	US-09-268-992-7	Sequence 7, App1
87	25	1.1	72604	3	US-09-657-474-7	Sequence 7, App1
88	25	1.1	74962	4	US-09-685-855A-3	Sequence 3, App1
89	25	1.1	80246	3	US-09-078-294-4	Sequence 4, App1
90	25	1.1	84995	3	US-09-797-906-3	Sequence 3, App1
91	25	1.1	107820	4	US-09-792-616-1	Sequence 1, App1
92	25	1.1	1107820	4	US-09-818-512-3	Sequence 3, App1
93	25	1.1	116592	4	US-09-818-512-3	Sequence 3, App1
94	25	1.1	116592	4	US-09-128-155-16	Sequence 16, App1
95	25	1.1	152331	3	US-09-128-155-17	Sequence 17, App1
96	25	1.1	176373	3	US-09-822-871-3	Sequence 3, App1
97	25	1.1	254566	4	US-09-539-333D-1	Sequence 1, App1
98	25	1.1	319608	4	US-09-539-333D-1	Sequence 1, App1
99	25	1.1	319608	4	US-09-679-409-1	Sequence 1, App1
100	25	1.1	319608	4	US-09-679-409-1	Sequence 1, App1

```
C 101 23 1.1 319608 4 US-09-679-409-1 Sequence 1, Appl1
C 102 23 1.1 392000 4 US-10-027-983-11 Sequence 11, Appl
C 103 22 1.1 92 4 US-09-513-999C-11042 Sequence 16042, A
C 104 22 1.1 186 4 US-09-513-999C-21364 Sequence 21364, A
C 105 22 1.1 195 4 US-09-513-999C-33224 Sequence 33224, A
C 106 22 1.1 195 4 US-09-513-999C-33365 Sequence 33365, A
C 107 22 1.1 216 4 US-09-513-999C-26690 Sequence 26690, A
C 108 22 1.1 388 2 US-08-967-101-68 Sequence 68, Appl
C 109 22 1.1 388 2 US-08-592-541-68 Sequence 68, Appl
C 110 22 1.1 388 3 US-09-124-698-68 Sequence 68, Appl
C 111 22 1.1 388 3 US-09-127-480-68 Sequence 68, Appl
C 112 22 1.1 388 3 US-08-996-841C-68 Sequence 68, Appl
C 113 22 1.1 388 3 US-09-124-523-68 Sequence 68, Appl
C 114 22 1.1 388 4 US-09-636-796A-68 Sequence 68, Appl
C 115 22 1.1 388 4 US-08-431-048F-68 Sequence 68, Appl
C 116 22 1.1 396 4 US-09-640-173-46 Sequence 46, Appl
C 117 22 1.1 396 4 US-09-713-550-46 Sequence 46, Appl
C 118 22 1.1 396 4 US-09-825-294-46 Sequence 46, Appl
C 119 22 1.1 396 4 US-09-970-966-46 Sequence 46, Appl
C 120 22 1.1 7720 3 US-09-513-999C-8781 Sequence 8781, Ap
C 121 22 1.1 10545 4 US-09-526-193A-14 Sequence 5, Appl1
C 122 22 1.1 10541 4 US-09-759-359A-3 Sequence 14, Appl
C 123 22 1.1 90541 4 US-10-207-973-3 Sequence 3, Appl1
C 124 22 1.1 106746 4 US-09-326-402C-1 Sequence 1, Appl1
C 125 22 1.1 106746 4 US-09-326-402C-12 Sequence 12, Appl
C 126 22 1.1 118067 4 US-09-497-855A-32 Sequence 32, Appl
C 127 22 1.1 221 4 US-09-513-999C-20878 Sequence 20878, A
C 128 21 1.0 275 4 US-09-513-999C-20385 Sequence 20385, A
C 129 21 1.0 851 4 US-10-204-708-65 Sequence 138, Appl
C 130 21 1.0 5981 4 US-09-854-133-441 Sequence 45, Appl
C 131 21 1.0 6020 4 US-10-204-708-8 Sequence 41, Appl
C 132 21 1.0 6566 4 US-10-204-708-75 Sequence 75, Appl
C 133 21 1.0 11050 4 US-10-204-708-85 Sequence 85, Appl
C 134 21 1.0 19233 4 US-10-204-708-45 Sequence 45, Appl
C 135 21 1.0 19513 4 US-10-204-708-39 Sequence 39, Appl
C 136 21 1.0 99500 3 US-09-798-096-10 Sequence 10, Appl
C 137 21 1.0 162450 3 US-09-345-882-1 Sequence 1, Appl1
C 138 21 1.0 786431 4 US-09-751-389-3 Sequence 3, Appl1
C 139 21 1.0 116 4 US-09-513-999C-18638 Sequence 18638, A
C 140 20 1.0 320 4 US-09-621-976-7941 Sequence 7941, Ap
C 141 20 1.0 330 4 US-09-513-999C-24340 Sequence 24340, A
C 142 20 1.0 336 4 US-09-513-999C-28590 Sequence 28590, A
C 143 20 1.0 387 4 US-09-621-976-8346 Sequence 8346, Ap
C 144 20 1.0 421 4 US-09-513-999C-35806 Sequence 35806, A
C 145 20 1.0 434 4 US-09-621-976-10150 Sequence 10150, A
C 146 20 1.0 464 4 US-09-621-976-18352 Sequence 18352, A
C 147 20 1.0 469 4 US-09-621-976-11355 Sequence 11355, A
C 148 20 1.0 474 4 US-09-621-976-13151 Sequence 13151, A
C 149 20 1.0 479 4 US-09-621-976-12796 Sequence 12796, A
C 150 20 1.0 2094 4 US-09-369-247-46 Sequence 46, Appl
C 151 20 1.0 3095 1 US-08-119-773-1 Sequence 1, Appl1
C 152 20 1.0 3095 4 US-09-814-915A-9 Sequence 9, Appl1
C 153 20 1.0 7676 1 US-08-451-778A-7 Sequence 7, Appl1
C 154 20 1.0 7676 2 US-08-451-778A-7 Sequence 7, Appl1
C 155 20 1.0 7676 2 US-08-998-208-7 Sequence 7, Appl1
C 156 20 1.0 7676 5 PCT-US95-06743-7 Sequence 7, Appl1
C 157 20 1.0 24707 4 US-09-740-027-3 Sequence 3, Appl1
C 158 20 1.0 30350 4 US-10-118-328-3 Sequence 3, Appl1
C 159 20 1.0 41684 4 US-09-536-059-1 Sequence 1, Appl1
C 160 20 1.0 63588 4 US-09-873-404-3 Sequence 3, Appl1
C 161 20 1.0 63588 4 US-10-243-735-3 Sequence 3, Appl1
C 162 20 1.0 81001 4 US-09-750-580-1 Sequence 1, Appl1
C 163 20 1.0 99916 4 US-09-816-095-3 Sequence 3, Appl1
C 164 20 1.0 118999 4 US-09-791-105B-32 Sequence 32, Appl1
C 165 20 1.0 90 4 US-09-513-999C-15370 Sequence 15370, A
C 166 19 0.9 69 4 US-09-513-999C-16037 Sequence 16037, A
C 167 19 0.9 163 4 US-09-513-999C-16023 Sequence 16023, A
C 168 19 0.9 193 4 US-09-513-999C-20284 Sequence 20284, A
C 169 19 0.9 328 4 US-09-513-999C-8772 Sequence 8772, Ap
C 170 19 0.9 423 4 US-09-513-999C-16646 Sequence 16646, A
C 171 19 0.9 454 3 US-08-991-789A-49 Sequence 49, Appl1
C 172 19 0.9 454 3 US-08-991-789A-49 Sequence 49, Appl1
C 173 19 0.9 454 3 US-08-991-789A-49 Sequence 49, Appl1
```

```
C 174 19 0.9 454 3 US-09-062-451-49 Sequence 49, Appl1
C 175 19 0.9 454 4 US-09-598-328-49 Sequence 49, Appl
C 176 19 0.9 454 4 US-09-289-198-49 Sequence 49, Appl
C 177 19 0.9 454 4 US-09-429-755-49 Sequence 49, Appl
C 178 19 0.9 597 4 US-09-252-991A-12149 Sequence 12149, A
C 179 19 0.9 1095 4 US-09-252-991A-12339 Sequence 12339, A
C 180 19 0.9 1302 1 US-08-525-697-1 Sequence 1, Appl1
C 181 19 0.9 1653 4 US-09-023-655-536 Sequence 536, Appl
C 182 19 0.9 1685 6 5340934-1 Patent No. 5340934
C 183 19 0.9 1740 4 US-09-252-991A-12081 Sequence 12081, A
C 184 19 0.9 11131 4 US-10-204-708-28 Sequence 28, Appl
C 185 19 0.9 14485 4 US-09-876-218-3 Sequence 3, Appl1
C 186 19 0.9 16595 3 US-09-146-053-7 Sequence 7, Appl1
C 187 19 0.9 19736 3 US-09-740-035-3 Sequence 3, Appl1
C 188 19 0.9 51719 4 US-09-918-686-2 Sequence 2, Appl1
C 189 19 0.9 51719 4 US-09-918-686-2 Sequence 2, Appl1
C 190 19 0.9 55298 4 US-09-491-356C-1 Sequence 1, Appl1
C 191 19 0.9 92139 4 US-09-918-686-1 Sequence 1, Appl1
C 192 19 0.9 92139 4 US-09-918-686-1 Sequence 1, Appl1
C 193 18 0.9 113 4 US-09-513-999C-17509 Sequence 17509, A
C 194 18 0.9 119 4 US-09-513-999C-19383 Sequence 19383, A
C 195 18 0.9 149 4 US-09-513-999C-36304 Sequence 36304, A
C 196 18 0.9 263 4 US-09-621-976-12054 Sequence 12054, A
C 197 18 0.9 287 4 US-09-513-999C-9694 Sequence 9694, Ap
C 198 18 0.9 299 4 US-09-621-976-7522 Sequence 7922, Ap
C 199 18 0.9 302 4 US-09-621-976-8545 Sequence 8545, Ap
C 200 18 0.9 305 4 US-09-621-976-11672 Sequence 11672, A
```

## ALIGNMENTS

```
RESULT 1
US-09-513-999C-16265
; Sequence 16265, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Ducaslet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 16265
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 54 _
; OTHER INFORMATION: r=a or g
US-09-513-999C-16265

Query Match 1.2%; Score 25; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAATAATTTAAATTTACTG 2036
Db 143 CTTACTAAATAATTTAAATTTACTG 167

RESULT 2
US-09-513-999C-24183
; Sequence 24183, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
```

```

; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 24183
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-24183

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 239;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 175 CTAATAAATAATTAATTAAGCTG 199

RESULT 3
US-09-470-276-64
; Sequence 64, Application US/09470276
; Patent No. 6670460
; GENERAL INFORMATION:
; APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.
; APPLICANT: KOLODNER, Richard
; APPLICANT: WINAND, Nena
; TITLE OF INVENTION: A METHOD OF DETECTION OF ALTERATIONS IN MSHS
; FILE REFERENCE: 700157/447483C
; CURRENT APPLICATION NUMBER: US/09/470,276
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/051,686
; PRIOR FILING DATE: 1997-07-03
; PRIOR APPLICATION NUMBER: PCT/US98/13850
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 64
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Human
; FEATURE: Intron
; NAME/KEY: Intron
; LOCATION: (412)..(413)
; OTHER INFORMATION: N = A or T or G or C
US-09-470-276-64

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 413;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 362 CTAATAAATAATTAATTAAGCTG 386

RESULT 4
US-09-621-976-14650
; Sequence 14650, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

```

```

; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14650
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14650

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 486;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 408 CTAATAAATAATTAATTAAGCTG 432

RESULT 5
US-09-621-976-3876
; Sequence 3876, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3876
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 23..232
US-09-621-976-3876

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 497;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 309 CTAATAAATAATTAATTAAGCTG 333

RESULT 6
US-09-621-976-1934/C
; Sequence 1934, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1934
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 1..291

```

```
/ NAME/KEY: sig peptide
/ LOCATION: 1..159
/ OTHER INFORMATION: Von Hejine matrix
/ OTHER INFORMATION: score 3.9000009536743
/ OTHER INFORMATION: seq SYLSTLYLHVSLT/IG
US-09-621-976-1934
```

```
Query Match          1.2%; Score 25; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
        |||||||
Db       348 CTACTAAAAATATATAAATTAGCTG 324
```

```
RESULT 7
US-09-183-266A-16
/ Sequence 16, Application US/09183266A
/ Patent No. 6361954
/ GENERAL INFORMATION:
/ APPLICANT: Stillman, Bruce
/ APPLICANT: Williams, R. Sanders
/ APPLICANT: Mendez, Juan
/ TITLE OF INVENTION: DNA REPLICATION-REGULATING GENES,
/ FILE REFERENCE: CSH:96-01A3
/ CURRENT APPLICATION NUMBER: US/09/183,266A
/ CURRENT FILING DATE: 1998-10-30
/ PRIOR APPLICATION NUMBER: PCT/US97/07333
/ PRIOR FILING DATE: 1997-05-02
/ PRIOR APPLICATION NUMBER: 08/648,650
/ PRIOR FILING DATE: 1996-05-15
/ PRIOR APPLICATION NUMBER: 08/643,034
/ PRIOR FILING DATE: 1996-05-02
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 16
/ LENGTH: 685
/ TYPE: DNA
/ ORGANISM: H. sapiens
US-09-183-266A-16
```

```
Query Match          1.2%; Score 25; DB 3; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
        |||||||
Db       303 CTACTAAAAATATATAAATTAGCTG 327
```

```
RESULT 8
US-09-641-638-629/C
/ Sequence 629, Application US/09641638
/ Patent No. 6432648
/ GENERAL INFORMATION:
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Cohen, Annick
/ TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
/ FILE REFERENCE: GENSET.051CPI
/ CURRENT APPLICATION NUMBER: US/09/641,638
/ CURRENT FILING DATE: 2000-08-16
/ PRIOR APPLICATION NUMBER: US 09/502,330
/ PRIOR FILING DATE: 2000-02-11
/ PRIOR APPLICATION NUMBER: US 60/133,200
/ PRIOR FILING DATE: 1999-05-07
/ PRIOR APPLICATION NUMBER: US 09/275,267
/ PRIOR FILING DATE: 1999-03-23
/ PRIOR APPLICATION NUMBER: US 60/119,917
```

```
/ PRIOR FILING DATE: 1999-02-12
/ NUMBER OF SEQ ID NOS: 1304
/ SOFTWARE: Patent.pm
/ SEQ ID NO 629
/ LENGTH: 1001
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 501
/ OTHER INFORMATION: 10-350-332 : polymorphic base C or T
/ NAME/KEY: misc_binding
/ LOCATION: 481..500
/ OTHER INFORMATION: 10-350-332.misl, potential
/ NAME/KEY: misc_binding
/ LOCATION: 502..521
/ OTHER INFORMATION: 10-350-332.mis2, potential complement
/ NAME/KEY: primer_bind
/ LOCATION: 172..189
/ OTHER INFORMATION: upstream amplification primer
/ NAME/KEY: primer_bind
/ LOCATION: 513..532
/ OTHER INFORMATION: downstream amplification primer, complement
/ NAME/KEY: misc_binding
/ LOCATION: 489..513
/ OTHER INFORMATION: 10-350-332 potential probe
US-09-641-638-629
```

```
Query Match          1.2%; Score 25; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
        |||||||
Db       733 CTACTAAAAATATATAAATTAGCTG 709
```

```
RESULT 9
US-09-641-638-630/C
/ Sequence 630, Application US/09641638
/ Patent No. 6432648
/ GENERAL INFORMATION:
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Cohen, Annick
/ TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
/ FILE REFERENCE: GENSET.051CPI
/ CURRENT APPLICATION NUMBER: US/09/641,638
/ CURRENT FILING DATE: 2000-08-16
/ PRIOR APPLICATION NUMBER: US 09/502,330
/ PRIOR FILING DATE: 2000-02-11
/ PRIOR APPLICATION NUMBER: US 60/133,200
/ PRIOR FILING DATE: 1999-05-07
/ PRIOR APPLICATION NUMBER: US 09/275,267
/ PRIOR FILING DATE: 1999-03-23
/ PRIOR APPLICATION NUMBER: US 60/119,917
/ PRIOR FILING DATE: 1999-02-12
/ NUMBER OF SEQ ID NOS: 1304
/ SOFTWARE: Patent.pm
/ SEQ ID NO 630
/ LENGTH: 1001
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 501
/ OTHER INFORMATION: 10-350-72 : polymorphic base C or T
/ NAME/KEY: misc_binding
/ LOCATION: 481..500
/ OTHER INFORMATION: 10-350-72.misl, potential
/ NAME/KEY: misc_binding
```



```
LOCATION: 502..521
OTHER INFORMATION: 10-350-72.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 430..447
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer_bind
LOCATION: 771..790
OTHER INFORMATION: downstream amplification primer, complement
NAME/KEY: misc_binding
LOCATION: 489..513
OTHER INFORMATION: 10-350-72 potential probe
US-09-641-638-630

Query Match      1.2%; Score 25; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAATATATAAATTAGCTG 2036
DB      991 CTACTAAATATATAAATTAGCTG 967

RESULT 10
US-10-170-097-629/c
Sequence 629, Application US/10170097
Patent No. 6794143
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GEN-T114XC2D1
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US/10/170,097
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 629
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
FEATURE:
NAME/KEY: misc_binding
LOCATION: 481..500
OTHER INFORMATION: 10-350-332.mis1, potential
FEATURE:
NAME/KEY: misc_binding
LOCATION: 502..521
OTHER INFORMATION: 10-350-332.mis2, potential complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 172..189
OTHER INFORMATION: upstream amplification primer
FEATURE:
NAME/KEY: primer_bind
LOCATION: 513..532
OTHER INFORMATION: downstream amplification primer, complement
FEATURE:
```

```
NAME/KEY: misc_binding
LOCATION: 489..513
OTHER INFORMATION: 10-350-332 potential probe
US-10-170-097-629

Query Match      1.2%; Score 25; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAATATATAAATTAGCTG 2036
DB      733 CTACTAAATATATAAATTAGCTG 709

RESULT 11
US-10-170-097-630/c
Sequence 630, Application US/10170097
Patent No. 6794143
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GEN-T114XC2D1
CURRENT FILING DATE: US/10/170,097
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 630
LENGTH: 1001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 501
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
FEATURE:
NAME/KEY: misc_binding
LOCATION: 481..500
OTHER INFORMATION: 10-350-72.mis1, potential
FEATURE:
NAME/KEY: misc_binding
LOCATION: 502..521
OTHER INFORMATION: 10-350-72.mis2, potential complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 430..447
OTHER INFORMATION: upstream amplification primer
FEATURE:
NAME/KEY: primer_bind
LOCATION: 771..790
OTHER INFORMATION: downstream amplification primer, complement
FEATURE:
NAME/KEY: misc_binding
LOCATION: 489..513
OTHER INFORMATION: 10-350-72 potential probe
US-10-170-097-630

Query Match      1.2%; Score 25; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2012 CTCTAAAAATATATAAATTAGCTG 2036  
DB 991 CTCTAAAAATATATAAATTAGCTG 967

RESULT 12  
US-09-148-545-105/c  
; Sequence 105, Application US/09148545  
; Patent No. 6590075  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 70 Human Secreted Proteins  
; FILE REFERENCE: P2001P1  
; CURRENT APPLICATION NUMBER: US/09/148,545  
; CURRENT FILING DATE: 1998-09-04  
; EARLIER APPLICATION NUMBER: PCT/US98/04482  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,161  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612

; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22

```
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 105
LENGTH: 1822
```

```
Query Match 1.2%; Score 25; DB 4; Length 1822;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTAAGTAAATATTAATTAAGCTG 2036
DB 180 CTAAGTAAATATTAATTAAGCTG 156
```

RESULT 13

```
US-09-232-200-46/c
Sequence 46; Application US/09232200A
Patent No. 6288213
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
FILE REFERENCE: WH197-21p3WB
CURRENT APPLICATION NUMBER: US/09/232,200A
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-07-20
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 46
LENGTH: 3694
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-200-46
```

```
Query Match 1.2%; Score 25; DB 3; Length 3694;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTAAGTAAATATTAATTAAGCTG 2036
DB 2580 CTAAGTAAATATTAATTAAGCTG 2556
```

```
RESULT 14
US-09-232-197-46/c
Sequence 46; Application US/09232197A
Patent No. 6300096
GENERAL INFORMATION:
APPLICANT: Stahl, Andreas
APPLICANT: Hirsch, David J.
APPLICANT: Lodish, Harvey F.
APPLICANT: Gimeno, Ruth E.
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
FILE REFERENCE: WH197-21p3MA
CURRENT APPLICATION NUMBER: US/09/232,197A
EARLIER APPLICATION NUMBER: 60/071,374
EARLIER FILING DATE: 1999-01-14
EARLIER APPLICATION NUMBER: 60/093,491
EARLIER FILING DATE: 1998-01-15
EARLIER APPLICATION NUMBER: 60/110,941
EARLIER FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 46
LENGTH: 3694
TYPE: DNA
ORGANISM: Homo sapiens
US-09-232-197-46
```

```
Query Match 1.2%; Score 25; DB 3; Length 3694;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTAAGTAAATATTAATTAAGCTG 2036
DB 2580 CTAAGTAAATATTAATTAAGCTG 2556
```

## RESULT 15

US-09-232-201-46/C  
; Sequence 46, Application US/09232201A  
; Patent No. 6348321  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WH197-21P3MC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; EARLIER FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 46  
; LENGTH: 3694  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-201-46

Query Match 1.2%; Score 25; DB 3; Length 3694;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
DB 2580 CTTACTAAAAATATATAAATTAGCTG 2556

## RESULT 16

US-09-232-195-46/C  
; Sequence 46, Application US/09232195A  
; Patent No. 6657049  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WH197-21P3MD  
; CURRENT APPLICATION NUMBER: US/09/232,195A  
; EARLIER FILING DATE: 1999-01-04  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 46  
; LENGTH: 3694  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-195-46

Query Match 1.2%; Score 25; DB 4; Length 3694;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
DB 2580 CTTACTAAAAATATATAAATTAGCTG 2556

## RESULT 17

US-09-232-200-24/C  
; Sequence 24, Application US/09232200A  
; Patent No. 6288213  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WH197-21P3MB  
; CURRENT APPLICATION NUMBER: US/09/232,200A  
; EARLIER FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 3704  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-200-24

Query Match 1.2%; Score 25; DB 3; Length 3704;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
DB 2580 CTTACTAAAAATATATAAATTAGCTG 2556

## RESULT 18

US-09-232-197-24/C  
; Sequence 24, Application US/09232197A  
; Patent No. 6300096  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WH197-21P3MA  
; CURRENT APPLICATION NUMBER: US/09/232,197A  
; EARLIER FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 3704  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-197-24

Query Match 1.2%; Score 25; DB 3; Length 3704;

Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 2580 CTACTAAAAATATATAAATTAGCTG 2556

## RESULT 19

US-09-232-201-24/C  
; Sequence 24, Application US/09232201A  
; Patent No. 6348321  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-213PMC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; EARLIER FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 3704  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (175)...(2112)  
US-09-232-201-24

Query Match 1.2%; Score 25; DB 3; Length 3704;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 2580 CTACTAAAAATATATAAATTAGCTG 2556

## RESULT 20

US-09-232-195-24/C  
; Sequence 24, Application US/09232195A  
; Patent No. 6657049  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-213MD  
; CURRENT APPLICATION NUMBER: US/09/232,195A  
; EARLIER FILING DATE: 1999-01-04  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 3704  
; TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: CDS  
LOCATION: (175)...(2112)  
US-09-232-195-24

Query Match 1.2%; Score 25; DB 4; Length 3704;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 2580 CTACTAAAAATATATAAATTAGCTG 2556

## RESULT 21

US-09-077-354B-3  
; Sequence 3, Application US/09077354B  
; Patent No. 6255096  
; GENERAL INFORMATION:  
; APPLICANT: HOPWOOD, JOHN JOSEPH; SCOTT, HAMISH STEELE;  
; APPLICANT: WEBER, BIRGIT; BLANCH, LIANNE; ANSON, DONALD STEWART  
; TITLE OF INVENTION: SYNTHETIC MAMMALIAN  
; TITLE OF INVENTION: '-N-ACETYLGLUCOSAMINIDASE AND GENETIC SEQUENCES ENCODING SAME  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 GARDEN CITY PLAZA  
; CITY: GARDEN CITY  
; STATE: NEW YORK  
; COUNTRY: UNITED STATES  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/077,354B  
; FILING DATE: 22-APRIL-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/00747  
; FILING DATE: 22-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FOKALSKY, ANN R.  
; REGISTRATION NUMBER: 34,697  
; REFERENCE/DOCKET NUMBER: 12416  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516 742 4343  
; TELEFAX: 516 742 4366  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10380 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: Chromosome 17  
; FEATURE:  
; NAME/KEY: exon 1  
; LOCATION: 990..1372  
; FEATURE:  
; NAME/KEY: exon 2  
; LOCATION: 2115..2262  
; FEATURE:  
; NAME/KEY: exon 3  
; LOCATION: 3056..3202  
; FEATURE:  
; NAME/KEY: exon 4  
; LOCATION: 3387..3472

FEATURE:  
NAME/KEY: exon 5  
LOCATION: 5667..5923  
FEATURE:  
NAME/KEY: exon 6  
LOCATION: 7745..8955  
US-09-077-354B-3

Query Match 1.2%; Score 25; DB 3; Length 10380;  
Best Local Similarity 100.0%; Pred.No.0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTTACTAAATATATAAATTAGCTG 2036  
Db 3672 CTTACTAAATATATAAATTAGCTG 3696

RESULT 22  
US-09-641-638-651/c  
Sequence 651, Application US/09641638  
Patent No. 6432648  
GENERAL INFORMATION:  
APPLICANT: Blumentfeld, Marta  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Chumakov, Ilya  
APPLICANT: Cohen, Annick  
TITLE OF INVENTION: BIOMOLECULAR MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM  
FILE REFERENCE: GENSET.051CPI  
CURRENT APPLICATION NUMBER: US/09/641,638  
PRIOR FILING DATE: 2000-08-16  
PRIOR APPLICATION NUMBER: US 09/502,330  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: US 60/133,200  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: US 09/275,267  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: US 60/119,917  
PRIOR FILING DATE: 1999-02-12  
NUMBER OF SEQ ID NOS: 1304  
SOFTWARE: Patent.pm  
SEQ ID NO 651  
LENGTH: 20674  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1123..3123  
OTHER INFORMATION: 5'regulatory region  
NAME/KEY: exon  
LOCATION: 3124..3297  
OTHER INFORMATION: exon 1  
NAME/KEY: exon  
LOCATION: 3871..4072  
OTHER INFORMATION: exon 2  
NAME/KEY: exon  
LOCATION: 5552..5633  
OTHER INFORMATION: exon 3  
NAME/KEY: exon  
LOCATION: 5758..5880  
OTHER INFORMATION: exon 4  
NAME/KEY: exon  
LOCATION: 5996..6099  
OTHER INFORMATION: exon 5  
NAME/KEY: exon  
LOCATION: 6349..6509  
OTHER INFORMATION: exon 6  
NAME/KEY: exon  
LOCATION: 7379..7522  
OTHER INFORMATION: exon 7  
NAME/KEY: exon  
LOCATION: 8645..8854  
OTHER INFORMATION: exon 8

NAME/KEY: exon  
LOCATION: 12254..12340  
OTHER INFORMATION: exon 9  
NAME/KEY: exon  
LOCATION: 12854..13023  
OTHER INFORMATION: exon 10  
NAME/KEY: exon  
LOCATION: 13308..13429  
OTHER INFORMATION: exon 11  
NAME/KEY: exon  
LOCATION: 16567..16667  
OTHER INFORMATION: exon 12  
NAME/KEY: exon  
LOCATION: 16775..16945  
OTHER INFORMATION: exon 13  
NAME/KEY: exon  
LOCATION: 17063..17554  
OTHER INFORMATION: exon 14  
NAME/KEY: misc\_feature  
LOCATION: 17555..20674  
OTHER INFORMATION: 3'regulatory region  
NAME/KEY: allele  
LOCATION: 1128  
OTHER INFORMATION: 10-508-191 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 1182  
OTHER INFORMATION: 10-508-245 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 1559  
OTHER INFORMATION: 10-509-284 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 1570  
OTHER INFORMATION: 10-509-295 : deletion of C  
NAME/KEY: allele  
LOCATION: 1627  
OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTT  
NAME/KEY: allele  
LOCATION: 2048  
OTHER INFORMATION: 10-511-62 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 2323  
OTHER INFORMATION: 10-511-337 : insertion of T  
NAME/KEY: allele  
LOCATION: 2341  
OTHER INFORMATION: 10-512-36 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 2623  
OTHER INFORMATION: 10-512-318 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 2832  
OTHER INFORMATION: 10-513-250 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 2844  
OTHER INFORMATION: 10-513-262 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 2934  
OTHER INFORMATION: 10-513-352 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 2947  
OTHER INFORMATION: 10-513-365 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 3802  
OTHER INFORMATION: 12-206-81 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 4062  
OTHER INFORMATION: 10-343-231 : deletion of C  
NAME/KEY: allele  
LOCATION: 4088  
OTHER INFORMATION: 12-206-366 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 4109  
OTHER INFORMATION: 10-343-278 : polymorphic base C or T  
NAME/KEY: allele

LOCATION: 4170  
OTHER INFORMATION: 10-343-339 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 5903  
OTHER INFORMATION: 10-346-23 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6019  
OTHER INFORMATION: 10-346-141 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6141  
OTHER INFORMATION: 10-346-263 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 6183  
OTHER INFORMATION: 10-346-305 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 6338  
OTHER INFORMATION: 10-347-74 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6375  
OTHER INFORMATION: 10-347-111 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 6429  
OTHER INFORMATION: 10-347-165 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 6467  
OTHER INFORMATION: 10-347-203 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6484  
OTHER INFORMATION: 10-347-220 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 6534  
OTHER INFORMATION: 10-347-271 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 6611  
OTHER INFORMATION: 10-347-348 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 7668  
OTHER INFORMATION: 10-348-391 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 8608  
OTHER INFORMATION: 10-349-47 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 8658  
OTHER INFORMATION: 10-349-97 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 8703  
OTHER INFORMATION: 10-349-142 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 8777  
OTHER INFORMATION: 10-349-216 : deletion of CTG  
NAME/KEY: allele  
LOCATION: 8785  
OTHER INFORMATION: 10-349-224 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 8826  
OTHER INFORMATION: 10-349-368 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 12171  
OTHER INFORMATION: 10-350-72 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 12429  
OTHER INFORMATION: 10-350-332 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 13341  
OTHER INFORMATION: 10-507-170 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 13492  
OTHER INFORMATION: 10-507-321 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 13524  
OTHER INFORMATION: 10-507-353 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 13535

Query Match. 1.2%; Score 25; DB 4; Length 20674;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTAATAAATAATAAATTAGCTG 2036  
DB 12661 CTAATAAATAATAAATTAGCTG 12637  
RESULT 23  
US-10-170-097-651/C  
; Sequence 651 Application US/10170097  
; Patent No. 6794143  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; FILE REFERENCE: GEN-T114XG2D1  
; CURRENT FILING DATE: US/10/170,097  
; PRIOR FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US 09/641,638  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 09/502,330  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: US 60/133,200  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: US 09/275,267  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: US 60/119,917  
; NUMBER OF SEQ ID NOS: 1304  
; SOFTWARE: Patent.pm  
; SEQ ID NO 651  
; LENGTH: 20674  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1123..3123  
; OTHER INFORMATION: 5' regulatory region  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 3124..3297  
; OTHER INFORMATION: exon 1  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 3871..4072  
; OTHER INFORMATION: exon 2  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 5552..5633  
; OTHER INFORMATION: exon 3  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 5758..5880  
; OTHER INFORMATION: exon 4  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 5996..6099  
; OTHER INFORMATION: exon 5  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 6349..6509  
; OTHER INFORMATION: exon 6  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 7379..7522  
; OTHER INFORMATION: exon 7  
; FEATURE:

```
/ NAME/KEY: exon
/ LOCATION: 8645..8854
/ OTHER INFORMATION: exon 8
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 12254..12340
/ OTHER INFORMATION: exon 9
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 12854..13023
/ OTHER INFORMATION: exon 10
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 13308..13429
/ OTHER INFORMATION: exon 11
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 16567..16667
/ OTHER INFORMATION: exon 12
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 16775..16945
/ OTHER INFORMATION: exon 13
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 17063..17554
/ OTHER INFORMATION: exon 14
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 17555..20674
/ OTHER INFORMATION: 3'regulatory region
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 1128
/ OTHER INFORMATION: 10-508-191 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 1182
/ OTHER INFORMATION: 10-508-245 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 1559
/ OTHER INFORMATION: 10-509-284 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 1570
/ OTHER INFORMATION: 10-509-295 : deletion of C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 1827
/ OTHER INFORMATION: 10-510-173 : variable motif ATTAA or TTTTTT
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 2048
/ OTHER INFORMATION: 10-511-62 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 2323
/ OTHER INFORMATION: 10-511-337 : insertion of T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 2341
/ OTHER INFORMATION: 10-512-36 : polymorphic base G or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 2623
/ OTHER INFORMATION: 10-512-318 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 2832
/ OTHER INFORMATION: 10-513-250 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
```

```
/ LOCATION: 2844
/ OTHER INFORMATION: 10-513-262 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 2934
/ OTHER INFORMATION: 10-513-352 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 2947
/ OTHER INFORMATION: 10-513-365 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 3802
/ OTHER INFORMATION: 12-206-81 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 4062
/ OTHER INFORMATION: 10-343-231 : deletion of C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 4088
/ OTHER INFORMATION: 12-206-366 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 4109
/ OTHER INFORMATION: 10-343-278 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 4170
/ OTHER INFORMATION: 10-343-339 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 5903
/ OTHER INFORMATION: 10-346-23 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6019
/ OTHER INFORMATION: 10-346-141 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6141
/ OTHER INFORMATION: 10-346-263 : polymorphic base G or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6183
/ OTHER INFORMATION: 10-346-305 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6338
/ OTHER INFORMATION: 10-347-74 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6375
/ OTHER INFORMATION: 10-347-111 : polymorphic base G or C
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6429
/ OTHER INFORMATION: 10-347-165 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6467
/ OTHER INFORMATION: 10-347-203 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 6484
/ OTHER INFORMATION: 10-347-220 : polymorphic base A or G
```

```
Query Match 1.2%? Score 25; DB 4; Length 20674;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2012 CTTACTAAAAATATAAAATTAGCTG 2036
|||||
```



DB 12661 CTACTAAATATATATATAGCTG 12637

RESULT 24  
US-09-776-7/c  
Sequence 7, Application US/09776976  
Patent No. 6566332

GENERAL INFORMATION:  
APPLICANT: Friebis, Joachim  
APPLICANT: Erickson, Mary Ruth  
APPLICANT: Yen, Frances  
APPLICANT: Bihain, Bernard  
TITLE OF INVENTION: OBG Globular Head and Uses Thereof for Decreasing Body Mass

FILE REFERENCE: 76.04.REG  
CURRENT APPLICATION NUMBER: US/09/776,976  
PRIOR FILING DATE: 2001-02-05  
PRIOR APPLICATION NUMBER: US 09/758,055  
PRIOR FILING DATE: 2001-01-10  
PRIOR APPLICATION NUMBER: US 60/176,228  
PRIOR FILING DATE: 2000-01-14  
PRIOR APPLICATION NUMBER: US 60/198,087  
PRIOR FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: US 60/299,881  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patent.pm  
SEQ ID NO 7

LENGTH: 20966  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
LOCATION: 1..4811  
OTHER INFORMATION: 5' regulatory region  
NAME/KEY: exon  
LOCATION: 4812..4851  
OTHER INFORMATION: exon 1  
NAME/KEY: exon  
LOCATION: 15144..15365  
OTHER INFORMATION: exon 2  
NAME/KEY: exon  
LOCATION: 16277..20559  
OTHER INFORMATION: exon 3  
NAME/KEY: misc.feature  
LOCATION: 20560..20966  
OTHER INFORMATION: 3' regulatory region  
NAME/KEY: allele  
LOCATION: 3787  
OTHER INFORMATION: 9-27-261 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 11118  
OTHER INFORMATION: 99-14387-129 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 15120  
OTHER INFORMATION: 9-12-48 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 15196  
OTHER INFORMATION: 9-12-124 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 15427  
OTHER INFORMATION: 9-12-355 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 15500  
OTHER INFORMATION: 9-12-428 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 15863  
OTHER INFORMATION: 99-14405-105 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 17170  
OTHER INFORMATION: 9-16-189 : polymorphic base deletion of A  
NAME/KEY: primer\_bind  
LOCATION: 3528..3545  
OTHER INFORMATION: 9-27.pu

NAME/KEY: primer\_bind  
LOCATION: 3928..3946  
OTHER INFORMATION: 9-27.tp complement  
NAME/KEY: primer\_bind  
LOCATION: 10990..11008  
OTHER INFORMATION: 99-14387.pu  
NAME/KEY: primer\_bind  
LOCATION: 11423..11442  
OTHER INFORMATION: 99-14387.tp complement  
NAME/KEY: primer\_bind  
LOCATION: 15073..15092  
OTHER INFORMATION: 9-12.pu  
NAME/KEY: primer\_bind  
LOCATION: 15503..15520  
OTHER INFORMATION: 9-12.tp complement  
NAME/KEY: primer\_bind  
LOCATION: 15759..15776  
OTHER INFORMATION: 99-14405.pu  
NAME/KEY: primer\_bind  
LOCATION: 16191..16211  
OTHER INFORMATION: 99-14405.tp complement  
NAME/KEY: primer\_bind  
LOCATION: 16982..17001  
OTHER INFORMATION: 9-16.pu  
NAME/KEY: primer\_bind  
LOCATION: 17384..17402  
OTHER INFORMATION: 9-16.tp complement  
NAME/KEY: misc\_binding  
LOCATION: 3775..3799  
OTHER INFORMATION: 9-27-261.probe  
NAME/KEY: misc\_binding  
LOCATION: 11106..11130  
OTHER INFORMATION: 99-14387-129.probe  
NAME/KEY: misc\_binding  
LOCATION: 15108..15132  
OTHER INFORMATION: 9-12-48.probe  
NAME/KEY: misc\_binding  
LOCATION: 15184..15208  
OTHER INFORMATION: 9-12-124.probe  
NAME/KEY: misc\_binding  
LOCATION: 15415..15439  
OTHER INFORMATION: 9-12-355.probe  
NAME/KEY: misc\_binding  
LOCATION: 15486..15512  
OTHER INFORMATION: 9-12-428.probe  
NAME/KEY: misc\_binding  
LOCATION: 15851..15875  
OTHER INFORMATION: 99-14405-105.probe  
NAME/KEY: misc\_binding  
LOCATION: 17158..17182  
OTHER INFORMATION: 9-16-189.probe  
NAME/KEY: primer\_bind  
LOCATION: 3768..3786  
OTHER INFORMATION: 9-27-261.mis  
NAME/KEY: primer\_bind  
LOCATION: 3788..3806  
OTHER INFORMATION: 9-27-261.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 11099..11117  
OTHER INFORMATION: 99-14387-129.mis  
NAME/KEY: primer\_bind  
LOCATION: 11119..11137  
OTHER INFORMATION: 99-14387-129.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15101..15119  
OTHER INFORMATION: 9-12-48.mis  
NAME/KEY: primer\_bind  
LOCATION: 15121..15139  
OTHER INFORMATION: 9-12-48.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15177..15195  
OTHER INFORMATION: 9-12-124.mis  
NAME/KEY: primer\_bind

```
/ LOCATION: 15197..15215
/ OTHER INFORMATION: 9-12-124.mis complement
/ NAME/KEY: primer bind
/ LOCATION: 15408..15426
/ OTHER INFORMATION: 9-12-355.mis
/ NAME/KEY: primer bind
/ LOCATION: 15428..15446
/ OTHER INFORMATION: 9-12-355.mis complement
/ NAME/KEY: primer bind
/ LOCATION: 15481..15499
/ OTHER INFORMATION: 9-12-428.mis
/ NAME/KEY: primer bind
/ LOCATION: 15501..15519
/ OTHER INFORMATION: 9-12-428.mis complement
/ NAME/KEY: primer bind
/ LOCATION: 15844..15862
/ OTHER INFORMATION: 99-14405-105.mis
/ NAME/KEY: primer bind
/ LOCATION: 15864..15882
/ OTHER INFORMATION: 99-14405-105.mis complement
/ NAME/KEY: primer bind
/ LOCATION: 17151..17169
/ OTHER INFORMATION: 9-16-189.mis
/ NAME/KEY: primer bind
/ LOCATION: 17171..17189
/ OTHER INFORMATION: 9-16-189.mis complement
/ US-09-776-976-7
```

```
Query Match 1.2%; Score 25; DB 4; Length 20966;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 6435 CTACTAAAAATATATAAATTAGCTG 6411
```

```
RESULT 25
US-09-909-547-7/c
/ Sequence 7, Application US/0909547
/ Patent No. 6579852
/ GENERAL INFORMATION:
/ APPLICANT: Friebis, Joachim
/ APPLICANT: Erickson, Mary Ruth
/ APPLICANT: Yen, Frances
/ APPLICANT: Bihain, Bernard
/ TITLE OF INVENTION: OB63 Globular Head and Uses Thereof for Decreasing Body Mass
/ FILE REFERENCE: 76.US6.CIP
/ CURRENT APPLICATION NUMBER: US/09/909,547
/ PRIOR FILING DATE: 2001-07-19
/ PRIOR APPLICATION NUMBER: US 09/776,976
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: US 09/758,055
/ PRIOR FILING DATE: 2001-01-10
/ PRIOR APPLICATION NUMBER: US 60/299,881
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: US 60/198,087
/ PRIOR FILING DATE: 2000-04-13
/ PRIOR APPLICATION NUMBER: US 60/176,228
/ PRIOR FILING DATE: 2000-01-14
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: Patent.pm
/ SEQ ID NO 7
/ LENGTH: 20966
/ TYPER: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 1..4811
/ OTHER INFORMATION: 5' regulatory region
/ NAME/KEY: exon
/ LOCATION: 4812..4851
/ OTHER INFORMATION: exon 1
```

```
/ NAME/KEY: exon
/ LOCATION: 15144..15365
/ OTHER INFORMATION: exon 2
/ NAME/KEY: exon
/ LOCATION: 16277..20559
/ OTHER INFORMATION: exon 3
/ NAME/KEY: misc feature
/ LOCATION: 20560..20966
/ OTHER INFORMATION: 3' regulatory region
/ NAME/KEY: allele
/ LOCATION: 3787
/ OTHER INFORMATION: 9-27-261 : polymorphic base G or C
/ NAME/KEY: allele
/ LOCATION: 1118
/ OTHER INFORMATION: 99-14387-129 : polymorphic base A or C
/ NAME/KEY: allele
/ LOCATION: 15120
/ OTHER INFORMATION: 9-12-48 : polymorphic base C or T
/ NAME/KEY: allele
/ LOCATION: 15196
/ OTHER INFORMATION: 9-12-124 : polymorphic base G or T
/ NAME/KEY: allele
/ LOCATION: 15427
/ OTHER INFORMATION: 9-12-355 : polymorphic base G or T
/ NAME/KEY: allele
/ LOCATION: 15500
/ OTHER INFORMATION: 9-12-428 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 15863
/ OTHER INFORMATION: 99-14405-105 : polymorphic base A or G
/ NAME/KEY: allele
/ LOCATION: 17170
/ OTHER INFORMATION: 9-16-189 : polymorphic base deletion of A
/ NAME/KEY: primer bind
/ LOCATION: 3528..3545
/ OTHER INFORMATION: 9-27.pu
/ NAME/KEY: primer bind
/ LOCATION: 3928..3946
/ OTHER INFORMATION: 9-27.rp complement
/ NAME/KEY: primer bind
/ LOCATION: 10990..11008
/ OTHER INFORMATION: 99-14387.pu
/ NAME/KEY: primer bind
/ LOCATION: 11423..11442
/ OTHER INFORMATION: 99-14387.rp complement
/ NAME/KEY: primer bind
/ LOCATION: 15073..15092
/ OTHER INFORMATION: 9-12.pu
/ NAME/KEY: primer bind
/ LOCATION: 15503..15520
/ OTHER INFORMATION: 9-12.rp complement
/ NAME/KEY: primer bind
/ LOCATION: 15759..15776
/ OTHER INFORMATION: 99-14405.pu
/ NAME/KEY: primer bind
/ LOCATION: 16191..16211
/ OTHER INFORMATION: 99-14405.rp complement
/ NAME/KEY: primer bind
/ LOCATION: 16982..17001
/ OTHER INFORMATION: 9-16.pu
/ NAME/KEY: primer bind
/ LOCATION: 17384..17402
/ OTHER INFORMATION: 9-16.rp complement
/ NAME/KEY: misc binding
/ LOCATION: 3775..3799
/ OTHER INFORMATION: 9-27-261.probe
/ NAME/KEY: misc binding
/ LOCATION: 11106..11130
/ OTHER INFORMATION: 99-14387-129.probe
/ NAME/KEY: misc binding
/ LOCATION: 15108..15132
/ OTHER INFORMATION: 9-12-48.probe
/ NAME/KEY: misc binding
```

LOCATION: 15184..15208  
OTHER INFORMATION: 9-12-124.probe  
NAME/KEY: misc\_binding  
LOCATION: 15415..15439  
OTHER INFORMATION: 9-12-355.probe  
NAME/KEY: misc\_binding  
LOCATION: 15488..15512  
OTHER INFORMATION: 9-12-428.probe  
NAME/KEY: misc\_binding  
LOCATION: 15851..15875  
OTHER INFORMATION: 99-14405-105.probe  
NAME/KEY: misc\_binding  
LOCATION: 17158..17182  
OTHER INFORMATION: 9-16-189.probe  
NAME/KEY: primer\_bind  
LOCATION: 3768..3786  
OTHER INFORMATION: 9-27-261.mis  
NAME/KEY: primer\_bind  
LOCATION: 3788..3806  
OTHER INFORMATION: 9-27-261.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 11099..11117  
OTHER INFORMATION: 99-14387-129.mis  
NAME/KEY: primer\_bind  
LOCATION: 11119..11137  
OTHER INFORMATION: 99-14387-129.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15101..15119  
OTHER INFORMATION: 9-12-48.mis  
NAME/KEY: primer\_bind  
LOCATION: 15121..15139  
OTHER INFORMATION: 9-12-48.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15177..15195  
OTHER INFORMATION: 9-12-124.mis  
NAME/KEY: primer\_bind  
LOCATION: 15197..15215  
OTHER INFORMATION: 9-12-124.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15408..15426  
OTHER INFORMATION: 9-12-355.mis  
NAME/KEY: primer\_bind  
LOCATION: 15428..15446  
OTHER INFORMATION: 9-12-355.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15481..15499  
OTHER INFORMATION: 9-12-428.mis  
NAME/KEY: primer\_bind  
LOCATION: 15501..15519  
OTHER INFORMATION: 9-12-428.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 15844..15862  
OTHER INFORMATION: 99-14405-105.mis  
NAME/KEY: primer\_bind  
LOCATION: 15864..15882  
OTHER INFORMATION: 99-14405-105.mis complement  
NAME/KEY: primer\_bind  
LOCATION: 17151..17169  
OTHER INFORMATION: 9-16-189.mis  
NAME/KEY: primer\_bind  
LOCATION: 17171..17189  
OTHER INFORMATION: 9-16-189.mis complement  
US-09-909-547-7

Query Match 1.2%; Score 25; DB 4; Length 20966;  
Best Local Similarity 100.0%; Pred.No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATAAATAAATTAGCTG 2036  
DB 6435 CTAATAAATAATAAATAAATTAGCTG 6411

RESULT 26  
US-09-569-852B-1/c  
Sequence ID Application US/09569852B  
Patent No. 6582909  
GENERAL INFORMATION:  
APPLICANT: Bouguetelerc, Lydie  
APPLICANT: Bihain, Bernard  
APPLICANT: Denison, Blake  
APPLICANT: Yen-Potin, Frances  
TITLE OF INVENTION: APM1 Biallelic Markers and Uses Thereof  
FILE REFERENCE: GEN-T113XC2  
CURRENT APPLICATION NUMBER: US/09/569,852B  
PRIOR FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: PCT/IB99/01858  
PRIOR FILING DATE: 1999-11-04  
PRIOR APPLICATION NUMBER: US 09/434,848  
PRIOR FILING DATE: 1999-11-04  
PRIOR APPLICATION NUMBER: US 60/119,593  
PRIOR FILING DATE: 1998-02-10  
PRIOR APPLICATION NUMBER: US 60/107,113  
PRIOR FILING DATE: 1998-11-04  
SOFTWARE: Patent version 3.1  
SEQ ID NO 1  
LENGTH: 20966  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(4811)  
OTHER INFORMATION: 5' regulatory region  
NAME/KEY: primer\_bind  
LOCATION: (14683)..(14701)  
OTHER INFORMATION: 17-34-860.mis  
NAME/KEY: primer\_bind  
LOCATION: (14703)..(14721)  
OTHER INFORMATION: 17-34-860.mis complement  
NAME/KEY: primer\_bind  
LOCATION: (14738)..(14756)  
OTHER INFORMATION: 17-34-915.mis  
NAME/KEY: primer\_bind  
LOCATION: (14758)..(14776)  
OTHER INFORMATION: 17-34-915.mis complement  
NAME/KEY: primer\_bind  
LOCATION: (14796)..(14814)  
OTHER INFORMATION: 17-35-71.mis  
NAME/KEY: primer\_bind  
LOCATION: (14816)..(14834)  
OTHER INFORMATION: 17-35-71.mis complement  
NAME/KEY: primer\_bind  
LOCATION: (15031)..(15049)  
OTHER INFORMATION: 17-35-306.mis  
NAME/KEY: primer\_bind  
LOCATION: (15051)..(15069)  
OTHER INFORMATION: 17-35-306.mis complement  
NAME/KEY: primer\_bind  
LOCATION: (15101)..(15119)  
OTHER INFORMATION: 9-12-48.mis  
NAME/KEY: primer\_bind  
LOCATION: (15121)..(15139)  
OTHER INFORMATION: 9-12-48.mis complement  
NAME/KEY: primer\_bind  
LOCATION: (15177)..(15195)  
OTHER INFORMATION: 9-12-124.mis  
NAME/KEY: primer\_bind  
LOCATION: (15197)..(15215)  
OTHER INFORMATION: 9-12-124.mis complement  
NAME/KEY: primer\_bind  
LOCATION: (15408)..(15426)  
OTHER INFORMATION: 9-12-355.mis  
NAME/KEY: primer\_bind  
LOCATION: (15428)..(15446)  
OTHER INFORMATION: 9-12-355.mis complement

```
NAME/KEY: primer_bind
LOCATION: (15481)..(15499)
OTHER INFORMATION: 9-12-428.mis
NAME/KEY: primer_bind
LOCATION: (15501)..(15519)
OTHER INFORMATION: 9-12-428.mis complement
NAME/KEY: primer_bind
LOCATION: (15661)..(15679)
OTHER INFORMATION: 17-36-47.mis
NAME/KEY: primer_bind
LOCATION: (15681)..(15699)
OTHER INFORMATION: 17-36-47.mis complement
NAME/KEY: primer_bind
LOCATION: (15771)..(15789)
OTHER INFORMATION: 17-36-120.mis
NAME/KEY: primer_bind
LOCATION: (15791)..(15809)
OTHER INFORMATION: 17-36-120.mis complement
NAME/KEY: primer_bind
LOCATION: (15844)..(15862)
OTHER INFORMATION: 99-14405-105.mis
NAME/KEY: primer_bind
LOCATION: (15864)..(15882)
OTHER INFORMATION: 99-14405-105.mis complement
NAME/KEY: primer_bind
LOCATION: (17151)..(17169)
OTHER INFORMATION: 9-16-189.mis
NAME/KEY: primer_bind
LOCATION: (17171)..(17189)
OTHER INFORMATION: 9-16-189.mis complement
NAME/KEY: primer_bind
LOCATION: (17810)..(17828)
OTHER INFORMATION: 17-37-629.mis
NAME/KEY: primer_bind
LOCATION: (17830)..(17848)
OTHER INFORMATION: 17-37-629.mis complement
NAME/KEY: primer_bind
LOCATION: (17992)..(18010)
OTHER INFORMATION: 17-37-811.mis
NAME/KEY: primer_bind
LOCATION: (18012)..(18030)
OTHER INFORMATION: 17-37-811.mis complement
NAME/KEY: primer_bind
LOCATION: (18470)..(18488)
OTHER INFORMATION: 17-38-349.mis
NAME/KEY: primer_bind
LOCATION: (18490)..(18508)
OTHER INFORMATION: 17-38-349.mis complement
NAME/KEY: primer_bind
LOCATION: (926)..(944)
OTHER INFORMATION: 17-30-216.mis
NAME/KEY: primer_bind
LOCATION: (946)..(964)
OTHER INFORMATION: 17-30-216.mis complement
NAME/KEY: primer_bind
LOCATION: (3719)..(3737)
OTHER INFORMATION: 9-27-211.mis
NAME/KEY: primer_bind
LOCATION: (3739)..(3757)
OTHER INFORMATION: 9-27-211.mis complement
NAME/KEY: primer_bind
LOCATION: (3754)..(3772)
OTHER INFORMATION: 9-27-246.mis
NAME/KEY: primer_bind
LOCATION: (3774)..(3792)
OTHER INFORMATION: 9-27-246.mis complement
NAME/KEY: primer_bind
LOCATION: (3768)..(3786)
OTHER INFORMATION: 9-27-261.mis
NAME/KEY: primer_bind
LOCATION: (3788)..(3806)
OTHER INFORMATION: 9-27-261.mis complement
NAME/KEY: primer_bind
```

```
LOCATION: (5076)..(5094)
OTHER INFORMATION: 17-31-298.mis
NAME/KEY: primer_bind
LOCATION: (5096)..(5114)
OTHER INFORMATION: 17-31-298.mis complement
NAME/KEY: primer_bind
LOCATION: (5191)..(5209)
OTHER INFORMATION: 17-31-413.mis
NAME/KEY: primer_bind
LOCATION: (5211)..(5229)
OTHER INFORMATION: 17-31-413.mis complement
NAME/KEY: primer_bind
LOCATION: (5364)..(5385)
OTHER INFORMATION: 17-31-413.mis complement
NAME/KEY: primer_bind
LOCATION: (10618)..(10636)
OTHER INFORMATION: 17-32-24.mis
NAME/KEY: primer_bind
LOCATION: (10638)..(10656)
OTHER INFORMATION: 17-32-24.mis complement
NAME/KEY: primer_bind
LOCATION: (11020)..(11038)
OTHER INFORMATION: 99-14387-50.mis
NAME/KEY: primer_bind
LOCATION: (11040)..(11058)
OTHER INFORMATION: 99-14387-50.mis complement
NAME/KEY: primer_bind
LOCATION: (11099)..(11117)
OTHER INFORMATION: 99-14387-129.mis
NAME/KEY: primer_bind
LOCATION: (11119)..(11137)
OTHER INFORMATION: 99-14387-129.mis complement
NAME/KEY: primer_bind
LOCATION: (11169)..(11187)
OTHER INFORMATION: 99-14387-199.mis
NAME/KEY: primer_bind
LOCATION: (11189)..(11207)
OTHER INFORMATION: 99-14387-199.mis complement
NAME/KEY: primer_bind
LOCATION: (13954)..(13972)
OTHER INFORMATION: 17-33-TGAGACT.mis
NAME/KEY: primer_bind
LOCATION: (13974)..(13992)
OTHER INFORMATION: 17-33-TGAGACT.mis complement
NAME/KEY: exon
LOCATION: (4812)..(4851)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (15144)..(15365)
OTHER INFORMATION:
NAME/KEY: exon
LOCATION: (16277)..(20559)
OTHER INFORMATION:
NAME/KEY: misc_feature
LOCATION: (20560)..(20966)
OTHER INFORMATION: 3' regulatory region

Query Match      1.2%  Score 25;  DB 4;  Length 20966;
Best Local Similarity 100.0%;  Pred. No. 0.08;
Matches 25;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      2012 CTACTAAAAATATAAAATTAGCTG 2036
      |||||
Db      6435 CTACTAAAAATATAAAATTAGCTG 6411

RESULT 27
US-09-544-398B-8/c
; Sequence 8, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
```

```

; APPLICANT: Becker, Robert R.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 33769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (33739), (33749), (33758)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknd
US-09-544-398B-8

Query Match
Best Local Similarity 1.2%; Score 25; DB 4; Length 33769;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036
DB 10891 CTTACTAAAAATATATAAAATTAGCTG 10867

RESULT 28
US-09-543-771-8/C
; Sequence 8, Application US/09543771
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/543,771
; PRIOR FILING DATE: 2000-04-05
; EARLIER APPLICATION NUMBER: US 09/229,319
; EARLIER FILING DATE: 1999-01-13
; EARLIER APPLICATION NUMBER: US 60/071,449
; EARLIER FILING DATE: 1998-01-13
; EARLIER APPLICATION NUMBER: US 60/105,511
; NUMBER OF SEQ ID NOS: 62
; SEQ ID NO 8
; LENGTH: 33769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (33739), (33749), (33758)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknd
US-09-543-771-8

Query Match
Best Local Similarity 1.2%; Score 25; DB 4; Length 33769;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036
DB 10891 CTTACTAAAAATATATAAAATTAGCTG 10867

RESULT 29
US-09-820-004-3
; Sequence 3, Application US/09820004
; Patent No. 6649385
```

```

; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001201
; CURRENT APPLICATION NUMBER: US/09/820,004
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 40090
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(40090)
; OTHER INFORMATION: n = A,T,C or G
US-09-820-004-3

Query Match
Best Local Similarity 1.2%; Score 25; DB 4; Length 40090;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036
DB 20627 CTTACTAAAAATATATAAAATTAGCTG 20651

RESULT 30
US-09-916-204-3/C
; Sequence 3, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 48763
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION:
US-09-916-204-3

Query Match
Best Local Similarity 1.2%; Score 25; DB 4; Length 48763;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036
DB 31864 CTTACTAAAAATATATAAAATTAGCTG 31840

RESULT 31
US-10-282-048-3/C
; Sequence 3, Application US/10282048
; Patent No. 6632948
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/282,048
; CURRENT FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
```

LENGTH: 48763  
TYPE: DNA  
ORGANISM: Human  
US-10-282-048-3

Query Match 1.2%; Score 25; DB 4; Length 48763;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 31864 CTACTAAAAATATATAAATTAGCTG 31840

RESULT 32  
US-09-813-817-3/C  
Sequence 3, Application US/09813817  
Patent No. 6340583  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001178  
CURRENT APPLICATION NUMBER: US/09/813,817  
CURRENT FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Human  
US-09-813-817-3

Query Match 1.2%; Score 25; DB 3; Length 59065;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 11824 CTACTAAAAATATATAAATTAGCTG 11800

RESULT 33  
US-09-978-197-3/C  
Sequence 3, Application US/09978197  
Patent No. 6403353  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001178DIV  
CURRENT APPLICATION NUMBER: US/09/978,197  
CURRENT FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: 09/813,817  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Human  
US-09-978-197-3

Query Match 1.2%; Score 25; DB 4; Length 59065;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 11824 CTACTAAAAATATATAAATTAGCTG 11800

RESULT 34  
US-10-135-696-3/C  
Sequence 3, Application US/10135696  
Patent No. 6740513  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001178DIV  
CURRENT APPLICATION NUMBER: US/10/135,696  
CURRENT FILING DATE: 2002-05-01  
PRIOR APPLICATION NUMBER: 09/813,817  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: 09/978,197  
PRIOR FILING DATE: 2001-10-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-135-696-3

Query Match 1.2%; Score 25; DB 4; Length 59065;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 11824 CTACTAAAAATATATAAATTAGCTG 11800

RESULT 35  
US-09-800-960-3/C  
Sequence 3, Application US/09800960  
Patent No. 6387677  
GENERAL INFORMATION:  
APPLICANT: YE, Jane et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001158  
CURRENT APPLICATION NUMBER: US/09/800,960  
CURRENT FILING DATE: 2001-03-08  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 62804  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(62804)  
OTHER INFORMATION: n = A,T,C or G  
US-09-800-960-3

Query Match 1.2%; Score 25; DB 3; Length 62804;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 43730 CTACTAAAAATATATAAATTAGCTG 43706

RESULT 36  
US-10-096-960-3/C  
Sequence 3, Application US/10096960  
Patent No. 6664085  
GENERAL INFORMATION:  
APPLICANT: YE, Jane et al.

```

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001158DIV
; CURRENT APPLICATION NUMBER: US/10/096,960
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 09/800,960
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A,T,C or G
US-10-096-960-3

Query Match          1.2%; Score 25; DB 4; Length 62804;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      43730 CTACTAAAAATATATAAATTAGCTG 43706

RESULT 37
US-09-984-890-3/c
; Sequence 3, Application US/099848890
; Patent No. 6452156
; GENERAL INFORMATION:
; APPLICANT: VAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 75395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(75395)
; OTHER INFORMATION: n = A,T,C or G
US-09-984-890-3

Query Match          1.2%; Score 25; DB 4; Length 75395;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      38456 CTACTAAAAATATATAAATTAGCTG 38432

RESULT 38
US-10-274-194-3/c
; Sequence 3, Application US/10274194
; Patent No. 6706511
; GENERAL INFORMATION:
; APPLICANT: VAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001306DIV
; CURRENT APPLICATION NUMBER: US/10/274,194
```

```

; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 75395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(75395)
; OTHER INFORMATION: n = A,T,C or G
US-10-274-194-3

Query Match          1.2%; Score 25; DB 4; Length 75395;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      38456 CTACTAAAAATATATAAATTAGCTG 38432

RESULT 39
US-10-148-806-3
; Sequence 3, Application US/10148806
; Patent No. 6762042
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148,806
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-148-806-3

Query Match          1.2%; Score 25; DB 4; Length 114793;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      37855 CTACTAAAAATATATAAATTAGCTG 37879

RESULT 40
US-10-148-806-3/c
; Sequence 3, Application US/10148806
; Patent No. 6762042
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148,806
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
```

```
; NUMBER OF SEQ ID NOS : 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-148-806-3
```

```
Query Match          1.2%; Score 25; DB 4; Length 114793;
Best Local Similarity 100.0%; Pred.No.0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAAATTAGCTG 2036
Db      100830 CTACTAAAAATATATAAAATTAGCTG 100806
```

```
RESULT 41
US-10-172-911-11
; Sequence 11, Application US/10172911
; Patent No. 6743909
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF PTPN12 EXPRESSION
; FILE REFERENCE: PTS-0016
; CURRENT APPLICATION NUMBER: US/10/172,911
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 123
; SEQ ID NO 11
; LENGTH: 137000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-172-911-11
```

```
Query Match          1.2%; Score 25; DB 4; Length 137000;
Best Local Similarity 100.0%; Pred.No.0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAAATTAGCTG 2036
Db      128823 CTACTAAAAATATATAAAATTAGCTG 128847
```

```
RESULT 42
US-09-801-876B-3
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3
```

```
Query Match          1.2%; Score 25; DB 4; Length 148567;
Best Local Similarity 100.0%; Pred.No.0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAAATTAGCTG 2036
Db      76478 CTACTAAAAATATATAAAATTAGCTG 76502
```

```
RESULT 43
US-10-254-869-3
; Sequence 3, Application US/10254869
; Patent No. 6653117
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: CL001160DIV
; CURRENT APPLICATION NUMBER: US/10/254,869
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3
```

```
Query Match          1.2%; Score 25; DB 4; Length 148567;
Best Local Similarity 100.0%; Pred.No.0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAAATTAGCTG 2036
Db      76478 CTACTAAAAATATATAAAATTAGCTG 76502
```

```
RESULT 44
US-09-804-471A-3/c
; Sequence 3, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3
```

```
Query Match          1.2%; Score 25; DB 4; Length 174493;
Best Local Similarity 100.0%; Pred.No.0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAAATTAGCTG 2036
Db      30756 CTACTAAAAATATATAAAATTAGCTG 30732
```

```
RESULT 45
US-10-238-709-3/c
```



```
; Sequence 3, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164D1V
; CURRENT APPLICATION NUMBER: US/10/238,709
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3
```

```
Query Match 1.2%; Score 25; DB 4; Length 174493;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 30756 CTACTAAAAATATATAAATTAGCTG 30732
```

```
RESULT 46
US-09-497-855A-37
; Sequence 37, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37
```

```
Query Match 1.2%; Score 25; DB 4; Length 193303;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 23296 CTACTAAAAATATATAAATTAGCTG 23320
```

```
RESULT 47
US-09-497-855A-44
; Sequence 44, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
```

```
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 44
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-44
```

```
Query Match 1.2%; Score 25; DB 4; Length 193303;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 23296 CTACTAAAAATATATAAATTAGCTG 23320
```

```
RESULT 48
US-09-734-674-3
; Sequence 3, Application US/09734674
; Patent No. 6498022
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001018
; CURRENT APPLICATION NUMBER: US/09/734,674
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 202001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(202001)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-674-3
```

```
Query Match 1.2%; Score 25; DB 4; Length 202001;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 179974 CTACTAAAAATATATAAATTAGCTG 179998
```

```
RESULT 49
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereeto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th floor
; CITY: San Francisco
; STATE: CA
```

COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Filts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20

Query Match 1.2% Score 25; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred.No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAATAAAATATATAAATTAGCTG 2036  
Db 45781 CTAATAAAATATATAAATTAGCTG 45805

RESULT 50  
US-08-724-394A-21  
Sequence 21, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Laufer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tauchihaishi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereo  
NUMBER OF SEQUENCES: 31  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Filts, Renee A.

REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-21

Query Match 1.2% Score 25; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred.No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAATAAAATATATAAATTAGCTG 2036  
Db 45781 CTAATAAAATATATAAATTAGCTG 45805

RESULT 51  
US-08-724-394A-22  
Sequence 22, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Laufer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tauchihaishi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereo  
NUMBER OF SEQUENCES: 31  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Filts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:

```

NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HUA-H. CONTIG"
US-08-724-394A-22

Query Match      1.2% Score 25; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAATATATAAATTAGCTG 2036
Db      45781 CTACTAAATATATAAATTAGCTG 45805

RESULT 52
US-10-027-983-11
Sequence 11, Application US/10027983
Patent No. 6617162
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Mark P. Roach
TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
FILE REFERENCE: RTS-0340
CURRENT APPLICATION NUMBER: US/10/027,983
CURRENT FILING DATE: 2001-12-18
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 11
LENGTH: 392000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 137740
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 137742
OTHER INFORMATION: unknown
NAME/KEY: misc feature
LOCATION: (138122)...(138221)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: unsure
LOCATION: 145507
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 151967
OTHER INFORMATION: unknown
NAME/KEY: misc feature
LOCATION: (151967)...(1542066)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: unsure
LOCATION: 154217
OTHER INFORMATION: unknown
NAME/KEY: misc feature
LOCATION: (164037)...(164136)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (174657)...(174756)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (186224)...(186323)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (195242)...(195341)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: unsure
LOCATION: 202703
OTHER INFORMATION: unknown
NAME/KEY: misc feature
LOCATION: (202711)...(202870)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (206246)...(215602)
OTHER INFORMATION: n = A,T,C or G

```

```

NAME/KEY: misc feature
LOCATION: (218126)...(218225)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (220360)...(220459)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (222717)...(222816)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (223981)...(224080)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (227467)...(227586)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (230157)...(230256)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (232299)...(232398)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (236552)...(236651)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: (238789)...(248788)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: exon
LOCATION: (118288)...(119101)
OTHER INFORMATION: exon 1C
NAME/KEY: exon:intron junction
LOCATION: (151129)...(151130)
OTHER INFORMATION: exon 5:intron 5
NAME/KEY: exon:intron junction
LOCATION: (299248)...(299249)
OTHER INFORMATION: exon 9:intron 9
NAME/KEY: exon:intron junction
LOCATION: (348578)...(348579)
OTHER INFORMATION: exon 10:intron 10
NAME/KEY: intron
LOCATION: (348579)...(381838)
OTHER INFORMATION: intron 10
NAME/KEY: intron:exon junction
LOCATION: (386185)...(386186)
OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

Query Match      1.2% Score 25; DB 4; Length 392000;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAATATATAAATTAGCTG 2036
Db      323647 CTACTAAATATATAAATTAGCTG 323671

RESULT 53
US-09-621-976-14785/C
Sequence 14785, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 14785
LENGTH: 240
TYPE: DNA

```

ORGANISM: Homo sapiens  
US-09-621-976-14785

Query Match 1.2%; Score 24; DB 4; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAATATATAAATTAGCTG 2036  
Db 228 TACTAAATATATAAATTAGCTG 205

RESULT 54  
US-09-621-976-14867/C

Sequence 14867, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 14867  
LENGTH: 444  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-14867

Query Match 1.2%; Score 24; DB 4; Length 444;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAATATATAAATTAGCTG 2036  
Db 228 TACTAAATATATAAATTAGCTG 205

RESULT 55  
US-09-621-976-1665/C

Sequence 1665, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 1665  
LENGTH: 485  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 228..401  
NAME/KEY: sig\_peptide  
LOCATION: 228..371  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 5.1999980926514  
OTHER INFORMATION: seq AQLMLFVEGSIC/TI  
US-09-621-976-1665

Query Match 1.2%; Score 24; DB 4; Length 485;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAATATATAAATTAGCTG 2036  
Db 228 TACTAAATATATAAATTAGCTG 205

RESULT 56  
US-09-621-976-1668/C

Sequence 1668, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 1668  
LENGTH: 491  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 228..401  
NAME/KEY: sig\_peptide  
LOCATION: 228..371  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 5.1999980926514  
OTHER INFORMATION: seq AQLMLFVEGSIC/TI  
NAME/KEY: misc\_feature  
LOCATION: 478  
OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-1668

Query Match 1.2%; Score 24; DB 4; Length 491;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAATATATAAATTAGCTG 2036  
Db 228 TACTAAATATATAAATTAGCTG 205

RESULT 57  
US-09-404-879A-24/C

Sequence 24, Application US/09404879A  
Patent No. 6468546  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C2  
CURRENT APPLICATION NUMBER: US/09/404,879A  
CURRENT FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 393  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 24  
LENGTH: 531  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(531)  
OTHER INFORMATION: n = A,T,C or G  
US-09-404-879A-24

Query Match 1.2%; Score 24; DB 4; Length 531;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATTAATAATTAGCTG 2036  
DB 200 TACTAAAAATATTAATAATTAGCTG 177

RESULT 58  
US-09-338-933-24/C  
; Sequence 24, Application US/09338933  
; Patent No. 6488931  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer Lynn  
; APPLICANT: King, Gordon E.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF  
; FILE REFERENCE: 210121.462C1  
; CURRENT APPLICATION NUMBER: US/09/338,933  
; CURRENT FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 312  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(531)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-338-933-24

Query Match 1.2%; Score 24; DB 4; Length 531;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATTAATAATTAGCTG 2036  
DB 200 TACTAAAAATATTAATAATTAGCTG 177

RESULT 59  
US-09-215-681-24/C  
; Sequence 24, Application US/09215681A  
; Patent No. 6528253  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Frudakis, Tony N.  
; APPLICANT: King, Gordon E.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS  
; TITLE OF INVENTION: OF OVARIAN CANCER  
; FILE REFERENCE: 210121.463  
; CURRENT APPLICATION NUMBER: US/09/215,681A  
; CURRENT FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 310  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(531)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-215-681-24

Query Match 1.2%; Score 24; DB 4; Length 531;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATTAATAATTAGCTG 2036  
DB 200 TACTAAAAATATTAATAATTAGCTG 177

RESULT 60  
US-09-216-003A-24/C  
; Sequence 24, Application US/09216003A  
; Patent No. 6670463  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Frudakis, Tony N.  
; APPLICANT: King, Gordon E.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: 210121.462  
; CURRENT APPLICATION NUMBER: US/09/216,003A  
; CURRENT FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 310  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (472)  
; OTHER INFORMATION: where n is a, c, g or t  
; NAME/KEY: modified\_base  
; LOCATION: (494)  
; OTHER INFORMATION: where n is a, c, g or t  
US-09-216-003A-24

Query Match 1.2%; Score 24; DB 4; Length 531;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATTAATAATTAGCTG 2036  
DB 200 TACTAAAAATATTAATAATTAGCTG 177

RESULT 61  
US-09-667-857-24/C  
; Sequence 24, Application US/09667857  
; Patent No. 6699664  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Fling, Steven P.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Fanger, Gary Richard  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darick  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.462C5  
; CURRENT APPLICATION NUMBER: US/09/667,857  
; CURRENT FILING DATE: 2000-09-20  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(531)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-667-857-24

Query Match 1.2%; Score 24; DB 4; Length 531;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATTAATAATTAGCTG 2036  
DB 200 TACTAAAAATATTAATAATTAGCTG 177

Db 200 TACTAAAAATATATAAATTAGCTG 177

RESULT 62

US-08-816-617A-1/c

/ Sequence 1, Application US/08816617A

/ Patent No. 6022741

/ GENERAL INFORMATION:

/ APPLICANT: Ting, Jenny P.-Y.

/ APPLICANT: Piskurich, Janet

/ TITLE OF INVENTION: No. 6022741el Regulatory Genetic DNA that

/ TITLE OF INVENTION: Regulates the Class II Transactivator (CITRA)

/ NUMBER OF SEQUENCES: 4

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Bell, Seltzer, Park & Gibson

/ STREET: 1211 East Morehead Street

/ CITY: Charlotte

/ STATE: No. 6022741th Carolina

/ COUNTRY: United States

/ ZIP: 28234

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/816,617A

/ FILING DATE:

/ CLASSIFICATION: 435

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Sibley, Kenneth D.

/ REGISTRATION NUMBER: 31,665

/ REFERENCE/DOCKET NUMBER: 5470-143

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 919-420-2200

/ TELEFAX: 919-881-3175

/ INFORMATION FOR SEQ ID NO: 1:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 6678 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: double

/ TOPOLOGY: linear

/ MOLECULE TYPE: DNA (genomic)

US-08-816-617A-1

Query Match 1.2%; Score 24; DB 3; Length 6678;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCT 2035

Db 1369 CTACTAAAAATATATAAATTAGCT 1346

RESULT 63

US-09-754-250-3

/ Sequence 3, Application US/09754250

/ Patent No. 6376225

/ GENERAL INFORMATION:

/ APPLICANT: Wei, Ming-Hui et al

/ TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE

/ TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

/ TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF

/ FILE REFERENCE: CL001063

/ CURRENT APPLICATION NUMBER: US/09/754,250

/ CURRENT FILING DATE: 2001-01-05

/ NUMBER OF SEQ ID NOS: 5

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 3

/ LENGTH: 111282

/ TYPE: DNA

/ ORGANISM: Human

/ FEATURE:

/ NAME/KEY: misc\_feature

/ LOCATION: (1) ... (111282)

/ OTHER INFORMATION: n = A,T,C or G

US-09-754-250-3

Query Match 1.2%; Score 24; DB 3; Length 111282;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATATAAATTAGCTG 2036

Db 37127 TACTAAAAATATATAAATTAGCTG 37150

RESULT 64

US-09-513-999C-26863/c

/ Sequence 26863, Application US/09513999C

/ Patent No. 6783961

/ GENERAL INFORMATION:

/ APPLICANT: Dumas Maline Edwards, J.B.

/ APPLICANT: Duclerc, A.

/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

/ FILE REFERENCE: 59. US2, REG

/ CURRENT APPLICATION NUMBER: US/09/513,999C

/ CURRENT FILING DATE: 2000-02-24

/ PRIOR APPLICATION NUMBER: US 60/122,487

/ PRIOR FILING DATE: 1999-02-26

/ NUMBER OF SEQ ID NOS: 36681

/ SOFTWARE: Patent.pm

/ SEQ ID NO 26863

/ LENGTH: 163

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-09-513-999C-26863

Query Match 1.1%; Score 23; DB 4; Length 163;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2013 TACTAAAAATATATAAATTAGCT 2035

Db 35 TACTAAAAATATATAAATTAGCT 13

RESULT 65

US-09-513-999C-18352/c

/ Sequence 18352, Application US/09513999C

/ Patent No. 6783961

/ GENERAL INFORMATION:

/ APPLICANT: Dumas Maline Edwards, J.B.

/ APPLICANT: Duclerc, A.

/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

/ FILE REFERENCE: 59. US2, REG

/ CURRENT APPLICATION NUMBER: US/09/513,999C

/ CURRENT FILING DATE: 2000-02-24

/ PRIOR APPLICATION NUMBER: US 60/122,487

/ PRIOR FILING DATE: 1999-02-26

/ NUMBER OF SEQ ID NOS: 36681

/ SOFTWARE: Patent.pm

/ SEQ ID NO 18352

/ LENGTH: 175

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: misc\_feature

/ LOCATION: 134

/ OTHER INFORMATION: w=a or t

/ FEATURE:

/ NAME/KEY: misc\_feature

LOCATION: 159  
OTHER INFORMATION: S-g or C  
US-09-513-999C-18352

Query Match 1.1%; Score 23; DB 4; Length 175;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATAATATAAATTAGC 2034  
DB 58 CTACTAAATAATATAAATTAGC 36

RESULT 66  
US-08-410-804-22/C  
Sequence 22, Application US/08410804  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive. Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/410,804  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/259,514  
FILING DATE: 14-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1389  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO. 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-410-804-22

Query Match 1.1%; Score 23; DB 1; Length 288;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATAATATAAATTAGC 2034  
DB 27 CTACTAAATAATATAAATTAGC 5

RESULT 67  
US-08-259-514-22/C  
Sequence 22, Application US/08259514  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive. Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,514  
FILING DATE: 14-JUN-1994

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive. Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,514  
FILING DATE: 14-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9954  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO. 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 288 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-259-514-22

Query Match 1.1%; Score 23; DB 1; Length 288;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATAATATAAATTAGC 2034  
DB 27 CTACTAAATAATATAAATTAGC 5

RESULT 68  
US-08-858-311-22/C  
Sequence 22, Application US/08858311  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
ATTORNEY/AGENT INFORMATION:  
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive. Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/858,311  
FILING DATE: 14-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/410,804  
FILING DATE: 27-MAR-1995  
APPLICATION NUMBER: US 08/259,514  
FILING DATE: 14-JUN-1994

```

; ATTORNEY/AGENT INFORMATION:
;   NAME: Campbell, Cathryn
;   REGISTRATION NUMBER: 31,815
;   REFERENCE/DOCKET NUMBER: P-LJ 1389
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 535-9001
;   TELEFAX: (619) 535-8949
;   INFORMATION FOR SEQ ID NO: 22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 288 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   US-08-858-311-22

Query Match      1.1%; Score 23; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      27 CTACTAAAAATATATAAATTAGC 5

RESULT 69
US-09-513-999C-9318/c
; Sequence 9318, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Duclert, A.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;   Patent No. 6783961
;   FILE REFERENCE: 59, US2, REG
;   CURRENT APPLICATION NUMBER: US/09/513,999C
;   CURRENT FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/122,487
;   PRIOR FILING DATE: 1999-02-26
;   NUMBER OF SEQ ID NOS: 36681
;   SOFTWARE: Patent.pm
;   SEQ ID NO 9318
;   LENGTH: 316
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   US-09-513-999C-9318

Query Match      1.1%; Score 23; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      207 CTACTAAAAATATATAAATTAGC 185

RESULT 70
US-09-621-976-14099
; Sequence 14099, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET, 054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
;   SEQ ID NO 14099
;   LENGTH: 382
```

```

;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 372..373
;   OTHER INFORMATION: n=a, g, c or t
;   US-09-621-976-14099

Query Match      1.1%; Score 23; DB 4; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      212 CTACTAAAAATATATAAATTAGC 234

RESULT 71
US-09-513-999C-31049/c
; Sequence 31049, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Duclert, A.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;   Patent No. 6783961
;   FILE REFERENCE: 59, US2, REG
;   CURRENT APPLICATION NUMBER: US/09/513,999C
;   CURRENT FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/122,487
;   PRIOR FILING DATE: 1999-02-26
;   NUMBER OF SEQ ID NOS: 36681
;   SOFTWARE: Patent.pm
;   SEQ ID NO 31049
;   LENGTH: 403
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 351
;   OTHER INFORMATION: m=a or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 384
;   OTHER INFORMATION: w=a or t
;   US-09-513-999C-31049

Query Match      1.1%; Score 23; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      83 CTACTAAAAATATATAAATTAGC 61

RESULT 72
US-09-621-976-12675/c
; Sequence 12675, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET, 054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
;   SEQ ID NO 12675
;   LENGTH: 444
```



TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-12675

Query Match 1.1%; Score 23; DB 4; Length 444;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGC 2034  
DB 251 CTACTAAAAATATATAAATTAGC 229

RESULT 73  
US-08-250-314-1/c  
Sequence 1, Application US/08250314  
Patent No. 5545550  
GENERAL INFORMATION:  
APPLICANT: Grossberg, Sidney E.  
APPLICANT: Kushnaryov, Vladimir M.  
APPLICANT: Cashdollar, Lester William  
APPLICANT: Carrigan, Donald R.  
TITLE OF INVENTION: HUMAN VIRUS ASSOCIATED WITH CHRONIC  
TITLE OF INVENTION: FATIGUE IMMUNE DEFICIENCY SYNDROME (CFIDS) AND ASSAY  
TITLE OF INVENTION: THEREFORE  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thad F. Kryshak, Quarles & Brady  
STREET: 411 East Wisconsin Avenue  
CITY: Milwaukee  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/250,314  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,452  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kryshak, Thad F.  
REGISTRATION NUMBER: 19,428  
REFERENCE/DOCKET NUMBER: 65-053-9075-8  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (414) 277-5774  
TELEFAX: (414) 277-5781  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: JHK virus  
US-08-250-314-1

Query Match 1.1%; Score 23; DB 1; Length 694;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2012 CTACTAAAAATATATAAATTAGC 2034  
DB 51 CTACTAAAAATATATAAATTAGC 29

RESULT 74  
US-08-708-107-1/c  
Sequence 1, Application US/08708107  
Patent No. 5837750  
GENERAL INFORMATION:  
APPLICANT: Grossberg, Sidney E.  
APPLICANT: Kushnaryov, Vladimir M.  
APPLICANT: Cashdollar, Lester William  
APPLICANT: Carrigan, Donald R.  
TITLE OF INVENTION: HUMAN VIRUS ASSOCIATED WITH CHRONIC  
TITLE OF INVENTION: FATIGUE IMMUNE DEFICIENCY SYNDROME (CFIDS) AND ASSAY  
TITLE OF INVENTION: THEREFORE  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thad F. Kryshak, Quarles & Brady  
STREET: 411 East Wisconsin Avenue  
CITY: Milwaukee  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/708,107  
FILING DATE: 08-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/891,452  
FILING DATE: 28-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Kryshak, Thad F.  
REGISTRATION NUMBER: 19,428  
REFERENCE/DOCKET NUMBER: 65-053-9075-8  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (414) 277-5774  
TELEFAX: (414) 277-5781  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: JHK virus  
US-08-708-107-1

Query Match 1.1%; Score 23; DB 1; Length 694;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGC 2034  
DB 51 CTACTAAAAATATATAAATTAGC 29

RESULT 75  
US-09-495-050A-13  
Sequence 13, Application US/09495050A  
Patent No. 6492505  
GENERAL INFORMATION:  
APPLICANT: Roopa, Reddy  
APPLICANT: Guejter, Karl, J.  
APPLICANT: Au-Young, Janice  
TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P  
FILE REFERENCE: PA-0013 US  
CURRENT APPLICATION NUMBER: US/09/495,050A

```

; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6492505 226815CT1
US-09-495-050A-13

Query Match          1.1%; Score 23; DB 4; Length 803;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
Db      651 CTACTAAAAATATATAAATTAGC 673

RESULT 76
US-09-671-317-150
; Sequence 150, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIOMIMETIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62 US3 CIP
; CURRENT APPLICATION NUMBER: US/09/671,317
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm
; SEQ ID NO 150
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-323-385 : polymorphic base T or C
; NAME/KEY: misc.binding
; LOCATION: 502..520
; OTHER INFORMATION: 12-323-385.mis1, complement
; NAME/KEY: misc.binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-323-385.mis2, potential
; NAME/KEY: primer.bind
; LOCATION: 868..886
; OTHER INFORMATION: upstream amplification primer, complement
; NAME/KEY: primer.bind
; LOCATION: 416..435
; OTHER INFORMATION: downstream amplification primer
; NAME/KEY: misc.binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-323-385 potential probe
; NAME/KEY: misc.feature
; LOCATION: 303,710
; OTHER INFORMATION: n=a, g, c or t
US-09-671-317-150
```

```

Query Match          1.1%; Score 23; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
Db      123 CTACTAAAAATATATAAATTAGC 145

RESULT 77
US-09-659-791A-10/c
; Sequence 10, Application US/09659791A
; Patent No. 6383808
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF CLUSTERIN EXPRESSION
; FILE REFERENCE: RTS-0156
; CURRENT APPLICATION NUMBER: US/09/659,791A
; CURRENT FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 8133
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-659-791A-10

Query Match          1.1%; Score 23; DB 3; Length 8133;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
Db      3926 CTACTAAAAATATATAAATTAGC 3904

RESULT 78
US-09-173-914-6/c
; Sequence 6, Application US/09173914
; Patent No. 6171857
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Eric
; TITLE OF INVENTION: A No. 6171857el Leucine Zipper, KAP-1 and
; TITLE OF INVENTION: Methods of Regulating DNA Dependent Protein Kinase Activity
; FILE REFERENCE: B0877/701/HK
; CURRENT APPLICATION NUMBER: US/09/173,914
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 60/064,557
; EARLIER FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 14636
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-173-914-6

Query Match          1.1%; Score 23; DB 3; Length 14636;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGC 2034
Db      11519 CTACTAAAAATATATAAATTAGC 11497

RESULT 79
US-09-844-634-17/c
; Sequence 17, Application US/09844634
; Patent No. 6410324
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
```

```

; APPLICANT: Andrew T. Walt
; TITLE OF INVENTION: ANTISENSE MODULATION OF TUMOR NECROSIS FACTOR RECEPTOR 2 EXPRESSION
; FILE REFERENCE: RTS-0216
; CURRENT APPLICATION NUMBER: US/09/844,634
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 174
; SEQ ID NO 17
; LENGTH: 15602
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-844-634-17

Query Match
Best Local Similarity 100.0%; Score 23; DB 4; Length 15602;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGC 2034
Db 1120 CTTACTAAAAATATATAAAATTAGC 1098

RESULT 80
US-09-901-151-3
; Sequence 3, Application US/09901151
; Patent No. 6677144
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001155-CIP
; CURRENT APPLICATION NUMBER: US/09/901,151
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 09/799,344
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18400
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(18400)
; OTHER INFORMATION: n = A,T,C or G
US-09-901-151-3

Query Match
Best Local Similarity 100.0%; Score 23; DB 4; Length 18400;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGC 2034
Db 6851 CTTACTAAAAATATATAAAATTAGC 6873

RESULT 81
US-09-819-989-3
; Sequence 3, Application US/09819989
; Patent No. 6482629
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001200
; CURRENT APPLICATION NUMBER: US/09/819,989
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19650
```

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(19650)
; OTHER INFORMATION: n = A,T,C or G
US-09-819-989-3

Query Match
Best Local Similarity 100.0%; Score 23; DB 4; Length 19650;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGC 2034
Db 11902 CTTACTAAAAATATATAAAATTAGC 11924

RESULT 82
US-10-273-992-3
; Sequence 3, Application US/10273992
; Patent No. 6664093
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001200D1V
; CURRENT APPLICATION NUMBER: US/10/273,992
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19650
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(19650)
; OTHER INFORMATION: n = A,T,C or G
US-10-273-992-3

Query Match
Best Local Similarity 100.0%; Score 23; DB 4; Length 19650;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGC 2034
Db 11902 CTTACTAAAAATATATAAAATTAGC 11924

RESULT 83
US-08-814-095-7
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/06/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 35060 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Cosmid including ACHB
DESCRIPTION: /desc = "Cosmid including ACHB
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 7q22
FEATURE:
NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHB Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHB"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHB"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHB"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 27005..27274
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHB"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 27255..28007
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHB"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: terminator
LOCATION: 27385..27387
FEATURE:
NAME/KEY: exon
LOCATION: 28008..28129

IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHB"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: complement (34528..34895)
OTHER INFORMATION: /function= "arsenite resistance
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: complement (34092..34358)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: complement (33779..33963)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: complement (33493..33591)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: complement (33297..33408)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: complement (32959..33094)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: complement (32569..32628)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: complement (32386..32468)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: complement (31894..32080)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: complement (31363..31534)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: complement (31131..31284)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: complement (30816..31011)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: complement (30470..30626)
OTHER INFORMATION: /gene= "AR"
```

```

; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
US-08-814-095-7

Query Match 1.1% Score 23; DB 3; Length 35060;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATAATTAATAATTAGC 2034
Db 7117 CTACTAAATAATTAATAATTAGC 7139

RESULT 84
US-09-740-041-3/c
; Sequence 3, Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 66804
; TYPE: DNA
; ORGANISM: Human
US-09-740-041-3

Query Match 1.1% Score 23; DB 4; Length 66804;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATAATTAATAATTAGC 2034
Db 31511 CTACTAAATAATTAATAATTAGC 31489

RESULT 85
US-09-851-896-3/c
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freiler
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT
; FILE REFERENCE: RTS-0220
; CURRENT APPLICATION NUMBER: US/09/851,896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match 1.1% Score 23; DB 4; Length 70000;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATAATTAATAATTAGC 2034
Db 30952 CTACTAAATAATTAATAATTAGC 30930

RESULT 86
US-09-268-992-7
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match 1.1% Score 23; DB 3; Length 72604;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATAATTAATAATTAGC 2034
Db 56992 CTACTAAATAATTAATAATTAGC 57014

RESULT 87
US-09-657-474-7
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
```

```
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7
```

```
Query Match          1.1%; Score 23; DB 3; Length 72604;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      56992 CTACTAAAAATATATAAATTAGC 57014
```

```
RESULT 88
US-09-685-853A-3/C
; Sequence 3, Application US/09685853A
; Patent No. 6479270
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C1000871
; CURRENT APPLICATION NUMBER: US/09/685,853A
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/182,194
; PRIOR FILING DATE: 2000-02-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 74962
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(74962)
; OTHER INFORMATION: n = A,T,C or G
US-09-685-853A-3
```

```
Query Match          1.1%; Score 23; DB 4; Length 74962;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      44565 CTACTAAAAATATATAAATTAGC 44543
```

```
RESULT 89
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4
```

```
Query Match          1.1%; Score 23; DB 3; Length 80246;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      63312 CTACTAAAAATATATAAATTAGC 63334
```

```
RESULT 90
US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEIN, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1001151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-906-3
```

```
Query Match          1.1%; Score 23; DB 3; Length 84495;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTACTAAAAATATATAAATTAGC 2034
DB      65914 CTACTAAAAATATATAAATTAGC 65936
```

```
RESULT 91
US-09-792-616-1
; Sequence 1, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PYE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRPc) causing
; TITLE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PYE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 107820
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "n" can be an A or a T or a G or a C
US-09-792-616-1
```

```
Query Match          1.1%; Score 23; DB 4; Length 107820;
Best Local Similarity 100.0%; Pred. No. 0.67;
```

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGC 2034

Db 61118 CTACTAAAAATATATAAATTAGC 61140

## RESULT 92

US-09-792-616-1/c  
; Sequence 1, Application US/09792616  
; Patent No. 6780587  
; GENERAL INFORMATION:  
; APPLICANT: PXE International, Inc.  
; APPLICANT: University of Hawaii  
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (WRP6) causing  
; FILE REFERENCE: PXE-001  
; CURRENT APPLICATION NUMBER: US/09/792,616  
; CURRENT FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 107820  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "n" can be an A or a T or a G or a C  
US-09-792-616-1

Query Match 1.1%; Score 23; DB 4; Length 107820;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGC 2034

Db 5222 CTACTAAAAATATATAAATTAGC 5200

## RESULT 93

US-09-818-512-3  
; Sequence 3, Application US/09818512  
; Patent No. 6537780  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCE: CL001192  
; CURRENT APPLICATION NUMBER: US/09/818,512  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 116592  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(116592)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-818-512-3

Query Match 1.1%; Score 23; DB 4; Length 116592;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGC 2034

Db 79683 CTACTAAAAATATATAAATTAGC 79705

## RESULT 94

US-09-818-512-3/c  
; Sequence 3, Application US/09818512  
; Patent No. 6537780  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCE: CL001192  
; CURRENT APPLICATION NUMBER: US/09/818,512  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 116592  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(116592)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-818-512-3

Query Match 1.1%; Score 23; DB 4; Length 116592;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGC 2034

Db 65278 CTACTAAAAATATATAAATTAGC 65256

## RESULT 95

US-09-128-155-16/c  
; Sequence 16, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 152331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(152331)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16

Query Match 1.1%; Score 23; DB 3; Length 152331;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGC 2034

Db 10901 CTACTAAAAATATATAAATTAGC 10879

## RESULT 96

US-09-128-155-17  
; Sequence 17, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:

```
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; EARLIER FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17
```

```
Query Match          1.1%; Score 23; DB 3; Length 176373;
Best Local Similarity 100.0%; Pred.No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTTACTAAAAATATATAAATTAGC 2034
Db      169007 CTTACTAAAAATATATAAATTAGC 169029
```

```
RESULT 97
US-09-822-871-3/C
; Sequence 3, Application US/09822871
; Patent No. 6723547
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: CL001219
; CURRENT APPLICATION NUMBER: US/09/822,871
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 254366
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(254366)
; OTHER INFORMATION: n = A,T,C or G
US-09-822-871-3
```

```
Query Match          1.1%; Score 23; DB 4; Length 254366;
Best Local Similarity 100.0%; Pred.No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2012 CTTACTAAAAATATATAAATTAGC 2034
Db      38516 CTTACTAAAAATATATAAATTAGC 38494
```

```
RESULT 98
US-09-539-333D-1
; Sequence 1, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marra
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
```

```
; APPLICANT: Bihain, Bernard
; APPLICANT: Escloux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET-047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5' regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29388..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 64666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 65854..67854
; OTHER INFORMATION: 3' regulatory region g35018 gene
; NAME/KEY: exon
; LOCATION: 94124..94964
```



```

OTHER INFORMATION: exon g35017
FEATURE:
NAME/KEY: exon
LOCATION: 201188..201234
OTHER INFORMATION: exon S g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216836..216915
OTHER INFORMATION: exon V g35030 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 213818..215818
OTHER INFORMATION: 3' regulatory region g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215941
OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215975
OTHER INFORMATION: exon Rbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216661..216952
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216661..217061
OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 217027..217061
OTHER INFORMATION: exon Q1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 229647..229742
OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 230408..230721
OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231272..231412
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231787..231880
OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231870..231879
OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 237406..237428
OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239807
OTHER INFORMATION: exon N2 complement g34872 gene

```

```

FEATURE:
NAME/KEY: exon
LOCATION: 239719..239853
OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240569
OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240824
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..241685
OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240800..240993
OTHER INFORMATION: exon M51 complement g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 241686..243685
OTHER INFORMATION: 5' regulatory region g34872 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 290652..292652
OTHER INFORMATION: 3' regulatory region g34665 gene
FEATURE:
NAME/KEY: exon
LOCATION: 292653..292841

Query Match 1.1%; Score 23; DB 4; Length 319608;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTTAAATATATAAATTAGC 2034
Db 43268 CTACTTAAATATATAAATTAGC 43290

RESULT 99
US-09-539-333D-1/c
Sequence 1, Application US/09539333D
Patent No. 6476208
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Bihain, Bernard
APPLICANT: Essioux, Laurent
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
FILE REFERENCE: GENSET.047AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
CURRENT FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US 60/126,903

```

;; PRIOR FILING DATE: 1999-03-30  
;; PRIOR APPLICATION NUMBER: US 60/131,971  
;; PRIOR FILING DATE: 1999-04-30  
;; PRIOR APPLICATION NUMBER: US 60/132,065  
;; PRIOR FILING DATE: 1999-04-30  
;; PRIOR APPLICATION NUMBER: US 60/143,928  
;; PRIOR FILING DATE: 1999-07-14  
;; PRIOR APPLICATION NUMBER: US 60/145,915  
;; PRIOR FILING DATE: 1999-07-27  
;; PRIOR APPLICATION NUMBER: US 60/146,453  
;; PRIOR FILING DATE: 1999-07-29  
;; PRIOR APPLICATION NUMBER: US 60/146,452  
;; PRIOR FILING DATE: 1999-07-29  
;; PRIOR APPLICATION NUMBER: US 60/162,288  
;; PRIOR FILING DATE: 1999-10-28  
;; PRIOR APPLICATION NUMBER: US 09/416,384  
;; PRIOR FILING DATE: 1999-10-12  
;; NUMBER OF SEQ ID NOS: 231  
;; SOFTWARE: Patent.pm  
;; SEQ ID NO 1  
;; LENGTH: 319608  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; LOCATION: 31..1107  
;; OTHER INFORMATION: 5'regulatory region g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 1108..1289  
;; OTHER INFORMATION: exon A g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 14877..14920  
;; OTHER INFORMATION: exon B g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 18778..18862  
;; OTHER INFORMATION: exon Bbis g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 25593..25740  
;; OTHER INFORMATION: exon C g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 29388..29502  
;; OTHER INFORMATION: exon D g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 29967..30282  
;; OTHER INFORMATION: exon E g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 64666..64812  
;; OTHER INFORMATION: exon F g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 65505..65853  
;; OTHER INFORMATION: exon G g35018 gene  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; LOCATION: 65854..67854  
;; OTHER INFORMATION: 3'regulatory region g35018 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 94124..94964  
;; OTHER INFORMATION: exon g35017  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 201188..201234  
;; OTHER INFORMATION: exon S g35030 gene  
;; FEATURE:  
;; NAME/KEY: exon

;; LOCATION: 214676..214793  
;; OTHER INFORMATION: exon T g35030 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 215702..215746  
;; OTHER INFORMATION: exon U g35030 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 216836..216915  
;; OTHER INFORMATION: exon V g35030 gene  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; LOCATION: 213818..215818  
;; OTHER INFORMATION: 3'regulatory region g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 215819..215941  
;; OTHER INFORMATION: exon R complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 215819..215975  
;; OTHER INFORMATION: exon Rbis complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 216661..216952  
;; OTHER INFORMATION: exon Qbis complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 216661..217061  
;; OTHER INFORMATION: exon Q complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 217027..217061  
;; OTHER INFORMATION: exon Q1 complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 229647..229742  
;; OTHER INFORMATION: exon X complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 230408..230721  
;; OTHER INFORMATION: exon P complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 231272..231412  
;; OTHER INFORMATION: exon Obis complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 231787..231880  
;; OTHER INFORMATION: exon O2 complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 231870..231879  
;; OTHER INFORMATION: exon O1 complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 234174..234321  
;; OTHER INFORMATION: exon O complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 237406..237428  
;; OTHER INFORMATION: exon Nbis complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 239719..239807  
;; OTHER INFORMATION: exon N2 complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 239719..239853  
;; OTHER INFORMATION: exon N complement g34872 gene  
;; FEATURE:  
;; NAME/KEY: exon  
;; LOCATION: 240528..240569

```
OTHER INFORMATION: exon M117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M52 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240824
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..241685
OTHER INFORMATION: exon M1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240800..240993
OTHER INFORMATION: exon M51 complement g34872 gene
FEATURE:
NAME/KEY: misc.feature
LOCATION: 241685..243685
OTHER INFORMATION: 5'regulatory region g34872 gene
FEATURE:
NAME/KEY: misc.feature
LOCATION: 290652..292652
OTHER INFORMATION: 3'regulatory region g34665 gene
FEATURE:
NAME/KEY: exon
LOCATION: 292653..292841

Query Match 1.1%; Score 23; DB 4; Length 319608;
Best Local Similarity: 100.0%; Pred. No. 0.67;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGC 2034
Db 263033 CTACTAAAAATATATAAATTAGC 263011

RESULT 100
US-09-679-409-1
Sequence 1, Application US/09679409
Patent No. 6555316
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bouguetel, Lydie
APPLICANT: Essiloux, Laurent
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS
FILE REFERENCE: 53 US15.CIP
CURRENT APPLICATION NUMBER: US/09/679,409
CURRENT FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 09/539,333
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 09/416,384
PRIOR FILING DATE: 1998-10-12
PRIOR APPLICATION NUMBER: 60/168,088
PRIOR FILING DATE: 1999-11-30
NUMBER OF SEQ ID NOS: 134
SOFTWARE: Patent.pm
SEQ ID NO 1
```

```
LENGTH: 319608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 199122..201122
OTHER INFORMATION: 5'regulatory region
NAME/KEY: exon
LOCATION: 201123..201334
OTHER INFORMATION: exon S
NAME/KEY: exon
LOCATION: 201123..201560
OTHER INFORMATION: exon S2
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U
NAME/KEY: exon
LOCATION: 216836..216994
OTHER INFORMATION: exon V
NAME/KEY: exon
LOCATION: 216836..217077
OTHER INFORMATION: exon V2
NAME/KEY: exon
LOCATION: 217671..217764
OTHER INFORMATION: exon V1
NAME/KEY: exon
LOCATION: 227655..227736
OTHER INFORMATION: exon V4
NAME/KEY: exon
LOCATION: 238715..238919
OTHER INFORMATION: exon V3
NAME/KEY: exon
LOCATION: 240440..240673
OTHER INFORMATION: exon W
NAME/KEY: exon
LOCATION: 240440..241153
OTHER INFORMATION: exon W2
NAME/KEY: exon
LOCATION: 241072..241291
OTHER INFORMATION: exon X
NAME/KEY: exon
LOCATION: 244353..244561
OTHER INFORMATION: exon Y
NAME/KEY: exon
LOCATION: 246273..247802
OTHER INFORMATION: exon Z
NAME/KEY: misc.feature
LOCATION: 247803..249803
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 8316
OTHER INFORMATION: 99-27943-150 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 21672
OTHER INFORMATION: 99-27935-193 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 65485
OTHER INFORMATION: 8-128-33 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 95396
OTHER INFORMATION: 99-31960-363 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 107281
OTHER INFORMATION: 99-24656-260 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 160640
OTHER INFORMATION: 99-24639-163 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 160876
OTHER INFORMATION: 99-24634-108 : polymorphic base A or T
```

NAME/KEY: allele  
LOCATION: 168974 : polymorphic base A or G  
OTHER INFORMATION: 99-7652-162 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 170810 : polymorphic base A or G  
OTHER INFORMATION: 99-16100-147 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 173358 : polymorphic base A or G  
OTHER INFORMATION: 99-5862-167 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 189957 : polymorphic base A or G  
OTHER INFORMATION: 99-5919-215 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 197163 : polymorphic base A or G  
OTHER INFORMATION: 99-24658-410 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 200778 : polymorphic base A or G  
OTHER INFORMATION: 8-303-235 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 202651 : polymorphic base A or G  
OTHER INFORMATION: 8-300-221 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 202679 : polymorphic base A or G  
OTHER INFORMATION: 8-300-193 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 203378 : polymorphic base A or T  
OTHER INFORMATION: 8-299-128 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 204138 : polymorphic base A or T  
OTHER INFORMATION: 8-296-213 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 204605 : polymorphic base C or T  
OTHER INFORMATION: 8-252-190 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 204934 : polymorphic base A or G  
OTHER INFORMATION: 99-24644-194 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 205206 : polymorphic base A or C  
OTHER INFORMATION: 8-295-248 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 205329 : polymorphic base C or T  
OTHER INFORMATION: 8-295-125 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 206064 : polymorphic base A or G  
OTHER INFORMATION: 8-293-130 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 206545 : polymorphic base A or G  
OTHER INFORMATION: 8-292-198 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 207313 : polymorphic base A or G  
OTHER INFORMATION: 8-251-322 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 208285 : polymorphic base A or G  
OTHER INFORMATION: 8-289-322 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 208960 : polymorphic base C or T  
OTHER INFORMATION: 8-287-249 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 209123 : polymorphic base A or T  
OTHER INFORMATION: 8-287-86 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 209631 : polymorphic base A or G  
OTHER INFORMATION: 8-285-319 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 210361 : polymorphic base G or C  
OTHER INFORMATION: 8-283-278 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 210463 : polymorphic base A or G  
OTHER INFORMATION: 8-283-176 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 210486 : polymorphic base G or C  
OTHER INFORMATION: 8-283-153 : polymorphic base G or C  
NAME/KEY: allele

LOCATION: 210583 : polymorphic base C or T  
OTHER INFORMATION: 8-283-56 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 210879 : polymorphic base G or C  
OTHER INFORMATION: 8-282-345 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 210964 : polymorphic base G or T  
OTHER INFORMATION: 8-282-260 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 210979 : polymorphic base A or C  
OTHER INFORMATION: 8-282-245 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 211050 : polymorphic base A or G  
OTHER INFORMATION: 8-282-174 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 211132 : polymorphic base A or T  
OTHER INFORMATION: 8-282-92 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 211247 : polymorphic base A or G  
OTHER INFORMATION: 8-281-367 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 211315 : polymorphic base A or G  
OTHER INFORMATION: 8-281-299 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 211366 : polymorphic base G or C  
OTHER INFORMATION: 8-281-248 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 212520 : polymorphic base A or C  
OTHER INFORMATION: 8-279-197 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 212821 : polymorphic base C or T  
OTHER INFORMATION: 8-278-289 : polymorphic base C or T  
NAME/KEY: allele

Query Match 1.1%; Score 23; DB 4; Length 319608;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATAAAATTAGC 2034  
DB 43268 CTTACTAAAAATATAAAATTAGC 43290

Search completed: January 4, 2005, 22:38:35  
Job time : 207 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 4, 2005, 20:59:22 ; Search time 1080 Seconds  
(without alignments)  
10648.008 Million cell updates/sec

Title: US-09-508-658b-1

Perfect score: 2036

Sequence: 1 agaccgggagagacggcg999.....aaaataataaattagctg 2036

Scoring table: **OLIGO2NUC\***  
Gapop 60.0 , Gapext 60.0

Searched: 4176236 seqs, 2824127955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269	13.2	1536	US-10-027-632-160833	Sequence 160833, A
2	269	13.2	1536	US-10-027-632-160833	Sequence 160833, A
3	178	8.7	556	US-10-029-386-549	Sequence 549, App
4	175	8.6	175	US-10-029-386-14254	Sequence 14254, A
5	150	7.4	530	US-10-029-386-1389	Sequence 1389, App
6	150	7.4	1948	US-10-027-632-97991	Sequence 97991, A
7	150	7.4	1948	US-10-027-632-97991	Sequence 97991, A
8	150	7.4	1948	US-10-027-632-97991	Sequence 97991, A
9	150	7.4	1948	US-10-027-632-97991	Sequence 97991, A
10	146	7.2	146	US-10-029-386-15091	Sequence 15091, A
11	120	5.9	862	US-10-027-632-163597	Sequence 163597, A
12	120	5.9	862	US-10-027-632-163597	Sequence 163597, A

13	60	2.9	60	US-09-908-975-11363	Sequence 11363, A
14	26	1.3	465	US-10-027-632-66360	Sequence 66360, A
15	26	1.3	465	US-10-027-632-67240	Sequence 67240, A
16	26	1.3	465	US-10-027-632-295913	Sequence 295913, A
17	26	1.3	465	US-10-027-632-66360	Sequence 66360, A
18	26	1.3	465	US-10-027-632-67240	Sequence 67240, A
19	26	1.3	465	US-10-027-632-295913	Sequence 295913, A
20	26	1.3	4329	US-10-417-375-154	Sequence 154, App
21	26	1.2	201	US-10-741-601-11354	Sequence 11354, A
22	25	1.2	201	US-10-741-601-11355	Sequence 11355, A
23	25	1.2	201	US-10-741-601-14743	Sequence 14743, A
24	25	1.2	213	US-09-764-891-6973	Sequence 6973, App
25	25	1.2	230	US-10-242-535A-36118	Sequence 36118, A
26	25	1.2	230	US-10-065-783A-36118	Sequence 36118, A
27	25	1.2	239	US-10-242-535A-39800	Sequence 39800, A
28	25	1.2	239	US-10-085-783A-39800	Sequence 39800, A
29	25	1.2	282	US-09-867-701-808	Sequence 808, App
30	25	1.2	292	US-10-040-739-673	Sequence 673, App
31	25	1.2	323	US-10-242-535A-9776	Sequence 9776, App
32	25	1.2	323	US-10-085-783A-9776	Sequence 9776, App
33	25	1.2	347	US-10-242-535A-14757	Sequence 14757, A
34	25	1.2	347	US-10-085-783A-14757	Sequence 14757, A
35	25	1.2	370	US-10-674-124A-23964	Sequence 23964, A
36	25	1.2	394	US-10-027-632-19134	Sequence 19134, A
37	25	1.2	394	US-10-027-632-19134	Sequence 19134, A
38	25	1.2	401	US-09-795-668-1061	Sequence 1061, App
39	25	1.2	401	US-09-795-668-1061	Sequence 1061, App
40	25	1.2	401	US-09-946-807-1061	Sequence 1061, App
41	25	1.2	403	US-10-357-930-57523	Sequence 57523, A
42	25	1.2	413	US-09-470-276-64	Sequence 64, App1
43	25	1.2	413	US-10-660-386-64	Sequence 64, App1
44	25	1.2	419	US-10-242-535A-39808	Sequence 39808, A
45	25	1.2	419	US-10-085-783A-39808	Sequence 39808, A
46	25	1.2	427	US-10-357-930-17263	Sequence 17263, A
47	25	1.2	430	US-10-027-632-283895	Sequence 283895, A
48	25	1.2	434	US-10-027-632-283895	Sequence 283895, A
49	25	1.2	434	US-09-918-995-34710	Sequence 34710, A
50	25	1.2	450	US-10-674-124A-15261	Sequence 15261, A
51	25	1.2	461	US-10-027-632-670729	Sequence 670729, A
52	25	1.2	461	US-10-027-632-670729	Sequence 670729, A
53	25	1.2	463	US-10-674-124A-15436	Sequence 15436, A
54	25	1.2	463	US-10-674-124A-1319	Sequence 1319, App
55	25	1.2	467	US-10-357-930-47086	Sequence 47086, A
56	25	1.2	470	US-10-674-124A-4488	Sequence 4488, App
57	25	1.2	470	US-09-867-701-6710	Sequence 6710, App
58	25	1.2	474	US-09-918-995-2669	Sequence 2669, App
59	25	1.2	474	US-10-242-535A-24474	Sequence 24474, A
60	25	1.2	474	US-10-085-783A-24474	Sequence 24474, A
61	25	1.2	486	US-09-918-995-15293	Sequence 15293, A
62	25	1.2	488	US-09-918-995-34441	Sequence 34441, A
63	25	1.2	497	US-10-027-632-33038	Sequence 33038, A
64	25	1.2	497	US-10-027-632-72533	Sequence 72533, A
65	25	1.2	497	US-10-027-632-13038	Sequence 13038, A
66	25	1.2	497	US-10-027-632-72533	Sequence 72533, A
67	25	1.2	503	US-10-240-425-252	Sequence 252, App
68	25	1.2	508	US-10-027-632-265243	Sequence 265243, A
69	25	1.2	508	US-10-027-632-265243	Sequence 265243, A
70	25	1.2	515	US-10-102-524-1199	Sequence 1199, App
71	25	1.2	526	US-10-027-632-289987	Sequence 289987, A
72	25	1.2	526	US-10-027-632-289987	Sequence 289987, A
73	25	1.2	536	US-10-027-632-270728	Sequence 270728, A
74	25	1.2	536	US-10-027-632-270728	Sequence 270728, A
75	25	1.2	546	US-10-027-632-242121	Sequence 242121, A
76	25	1.2	546	US-10-027-632-242121	Sequence 242121, A
77	25	1.2	549	US-10-027-632-86816	Sequence 86816, A
78	25	1.2	549	US-10-027-632-86816	Sequence 86816, A
79	25	1.2	549	US-10-027-632-116486	Sequence 116486, A
80	25	1.2	549	US-10-027-632-116486	Sequence 116486, A
81	25	1.2	549	US-10-027-632-86815	Sequence 86815, A
82	25	1.2	549	US-10-027-632-86816	Sequence 86816, A
83	25	1.2	549	US-10-027-632-116486	Sequence 116486, A
84	25	1.2	552	US-10-027-632-116486	Sequence 116486, A
85	25	1.2	552	US-10-027-632-283821	Sequence 283821, A





```

; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14254
; LENGTH: 175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21 74.0
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: O43918, EVALUATE 7.00e-10
; OTHER INFORMATION: EST HUMAN HIT: B1040873.1, EVALUATE 1.50e-02
; OTHER INFORMATION: NT HIT: Z97990.1, EVALUATE 5.00e-94
; US-10-029-386-14254

Query Match      8.6%; Score 175; DB 15; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.7e-77;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      269 GAGAGCCTTCATCTGAAGAAAAGAGGGCTGCCCCAGGCCCTTCACAGCCCTCTGCTGCC 328
DB      1  GAGACCTTCATCTGAAGAAAAGAGGGCTGCCCCAGGCCCTTCACAGCCCTCTGCTGCC 60
QY      329 TGGCTGCTACCCAGAGCTCCAGAGCCTCTGAGCTTCTGAGGGGTGCTGTTCAGAGAC 388
DB      61  TGGCTGCTACCCAGAGCTCCAGAGCCTCTGAGCTTCTGAGGGGTGCTGTTCAGAGAC 120
QY      389 TACAACCTGGAGGCTATGGCCGGCTGACAGCCCATCTTGAGAGAGCTTCCCAAG 443
DB      121 TACAACCTGGAGGCTATGGCCGGCTGACAGCCCATCTTGAGAGAGCTTCCCAAG 175

RESULT 5
US-10-029-386-1389
; Sequence 1389, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1389
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21 74.0
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: O43918, EVALUATE 2.00e-22
; OTHER INFORMATION: NT HIT: AB006684.1, EVALUATE 0.00e-00
; OTHER INFORMATION: EST_HUMAN HIT: BG330062.1, EVALUATE 1.90e-01
; US-10-029-386-1389
```

```

Query Match      7.4%; Score 150; DB 15; Length 530;
Best Local Similarity 100.0%; Pred. No. 5.8e-65;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      786 CAGCGGCTCCAAAGATGATCCAGTTGGCGGAGATTCTTACACTCCAGCAAGTTCC 845
DB      103 CAGCGGCTCCAAAGATGATCCAGTTGGCGGAGATTCTTACACTCCAGCAAGTTCC 162
QY      846 AAGACTCCGCGCAGTGGGAAAGCAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTTGTTCC 905
DB      163 AAGACTCCGCGCAGTGGGAAAGCAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTTGTTCC 222
QY      906 GAGCCAGGAGGCCCGAGGCGCTGCCCCG 935
DB      223 GAGCCAGGAGGCCCGAGGCGCTGCCCCG 252

RESULT 6
US-10-027-632-97991
; Sequence 97991, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97991
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-97991

Query Match      7.4%; Score 150; DB 13; Length 1948;
Best Local Similarity 100.0%; Pred. No. 4.8e-65;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      786 CAGCGGCTCCAAAGATGATCCAGTTGGCGGAGATTCTTACACTCCAGCAAGTTCC 845
DB      1220 CAGCGGCTCCAAAGATGATCCAGTTGGCGGAGATTCTTACACTCCAGCAAGTTCC 1279
QY      846 AAGACTCCGCGCAGTGGGAAAGCAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTTGTTCC 905
DB      1280 AAGACTCCGCGCAGTGGGAAAGCAAGGCCCGCAGCAGCAGTGGCCCGAAGCCTTGTTCC 1339
QY      906 GAGCCAGGAGGCCCGAGGCGCTGCCCCG 935
DB      1340 GAGCCAGGAGGCCCGAGGCGCTGCCCCG 1369

RESULT 7
US-10-027-632-97992
; Sequence 97992, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```



```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 97991
; LENGTH: 1948
; ORGANISM: Human
; US-10-027-632-97992

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-65;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 CAGGCGGCTCCAAAGATGATCCAGTTGGCGGGAGTTCTACACTCCAGCAAGTTG 845
DB 1220 CAGGCGGCTCCAAAGATGATCCAGTTGGCGGGAGTTCTACACTCCAGCAAGTTG 1279
QY 846 AAGACTCCGGCAGTGGGAAAGCAAGCCCGCAGCAGCAGTGGCCGGAAGCTCTGTTTC 905
DB 1280 AAGACTCCGGCAGTGGGAAAGCAAGCCCGCAGCAGCAGTGGCCGGAAGCTCTGTTTC 1339
QY 906 GAGCCAAGGAGAGCCGAGGCGCTGCCCCG 935
DB 1340 GAGCCAAGGAGAGCCGAGGCGCTGCCCCG 1369

RESULT 8
US-10-027-632-97991
; Sequence 97991, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 97991
; LENGTH: 1948
; ORGANISM: Human
; US-10-027-632-97992

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-65;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 CAGGCGGCTCCAAAGATGATCCAGTTGGCGGGAGTTCTACACTCCAGCAAGTTG 845
DB 1220 CAGGCGGCTCCAAAGATGATCCAGTTGGCGGGAGTTCTACACTCCAGCAAGTTG 1279
QY 846 AAGACTCCGGCAGTGGGAAAGCAAGCCCGCAGCAGCAGTGGCCGGAAGCTCTGTTTC 905
DB 1280 AAGACTCCGGCAGTGGGAAAGCAAGCCCGCAGCAGCAGTGGCCGGAAGCTCTGTTTC 1339
QY 906 GAGCCAAGGAGAGCCGAGGCGCTGCCCCG 935
DB 1340 GAGCCAAGGAGAGCCGAGGCGCTGCCCCG 1369

RESULT 9
US-10-027-632-97992
; Sequence 97992, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 97992
; LENGTH: 1948
; ORGANISM: Human
; US-10-027-632-97992

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-65;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 CAGGCGGCTCCAAAGATGATCCAGTTGGCGGGAGTTCTACACTCCAGCAAGTTG 845
DB 1220 CAGGCGGCTCCAAAGATGATCCAGTTGGCGGGAGTTCTACACTCCAGCAAGTTG 1279
QY 846 AAGACTCCGGCAGTGGGAAAGCAAGCCCGCAGCAGCAGTGGCCGGAAGCTCTGTTTC 905
DB 1280 AAGACTCCGGCAGTGGGAAAGCAAGCCCGCAGCAGCAGTGGCCGGAAGCTCTGTTTC 1339
QY 906 GAGCCAAGGAGAGCCGAGGCGCTGCCCCG 935
DB 1340 GAGCCAAGGAGAGCCGAGGCGCTGCCCCG 1369

RESULT 10
US-10-029-386-15091
```

```
/ Sequence 15091, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 15091
/ LENGTH: 146
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR21, 74.0
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.94
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
/ OTHER INFORMATION: EST HUMAN HIT: B6330062.1, EVALUE 4.70e-02
/ OTHER INFORMATION: SWISSPROT HIT: O43918, EVALUE 2.00e-22
/ OTHER INFORMATION: NT HIT: AB006684.1, EVALUE 8.00e-77
US-10-029-366-15091
```

```
Query Match
Best Local Similarity 7.2%; Score 146; DB 15; Length 146;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 789 GCGGCTCCAGAGTCATCGAGTGGCGGGAGTTCTACATCCAGCAATTCGAG 848
DB 1 GCGGCTCCAGAGTCATCGAGTGGCGGGAGTTCTACATCCAGCAATTCGAG 60
QY 849 ACTCCGCGAGTGGAGAACAGAGCCGCGAGCAGCAGTGGCCGGAAGCTTGGTTCGAG 908
DB 61 ACTCCGCGAGTGGAGAACAGAGCCGCGAGCAGCAGTGGCCGGAAGCTTGGTTCGAG 120
QY 909 CCAAGGAGCCCGAGGGGCGCTGCCCCC 934
DB 121 CCAAGGAGCCCGAGGGGCGCTGCCCCC 146
```

## RESULT 11

```
US-10-027-632-163597/c
/ Sequence 163597, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
```

```
/ SEQ ID NO 163597
/ LENGTH: 862
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-163597
```

```
Query Match
Best Local Similarity 5.9%; Score 120; DB 13; Length 862;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1012 CCAGAGATGAGAGCAGTGTGCGGTGTGTCGGACGGCGGAGCTCATCTGTCGGA 1071
DB 329 CCAGAGATGAGAGCAGTGTGCGGTGTGTCGGACGGCGGAGCTCATCTGTCGGA 270
QY 1072 GCGCTGCCCTCGGGGCTTCCACCTGCGCTGCTGCCCTCCGCTCCGGAGATCCCGAG 1131
DB 269 GCGCTGCCCTCGGGGCTTCCACCTGCGCTGCTGCCCTCCGCTCCGGAGATCCCGAG 210
```

## RESULT 12

```
US-10-027-632-163597/c
/ Sequence 163597, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 163597
/ LENGTH: 862
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-163597
```

```
Query Match
Best Local Similarity 5.9%; Score 120; DB 15; Length 862;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1012 CCAGAGATGAGAGCAGTGTGCGGTGTGTCGGACGGCGGAGCTCATCTGTCGGA 1071
DB 329 CCAGAGATGAGAGCAGTGTGCGGTGTGTCGGACGGCGGAGCTCATCTGTCGGA 270
QY 1072 GCGCTGCCCTCGGGGCTTCCACCTGCGCTGCTGCCCTCCGCTCCGGAGATCCCGAG 1131
DB 269 GCGCTGCCCTCGGGGCTTCCACCTGCGCTGCTGCCCTCCGCTCCGGAGATCCCGAG 210
```

## RESULT 13

```
US-09-908-975-11363
/ Sequence 11363, Application US/09908975
/ Publication No. US20030165843A1
/ GENERAL INFORMATION:
/ APPLICANT: SHOSHAN, Avi
/ APPLICANT: WASSERMAN, Alon
/ APPLICANT: MINTZ, Eli
```

```
APPLICANT: MINTZ, Liat
APPLICANT: RAIGER, Simon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICER
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 11363
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-09-908-975-11363
```

```
Query Match
Best Local Similarity 100.0%; Score 60; DB 10; Length 60;
Pred. No. 2,1e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1894 GCTCGGCTGTAAACAGCTCTGTCTTCTGGGACACGACGATCATGTGCTGGAAATTA 1953
Db 1 GCTCGGCTGTAAACAGCTCTGTCTTCTGGGACACGACGATCATGTGCTGGAAATTA 60
```

```
RESULT 14
US-10-027-632-66360/c
Sequence 66360, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 66360
LENGTH: 465
TYPE: DNA
ORGANISM: Human
US-10-027-632-66360
```

```
Query Match
Best Local Similarity 100.0%; Score 26; DB 13; Length 465;
Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2011 CCTACTAAAAATATATAAATTAGCTG 2036
Db 65 CCTACTAAAAATATATAAATTAGCTG 40
```

```
RESULT 15
US-10-027-632-67240/c
Sequence 67240, Application US/10027632
```

```
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 67240
LENGTH: 465
TYPE: DNA
ORGANISM: Human
US-10-027-632-67240
```

```
Query Match
Best Local Similarity 100.0%; Score 26; DB 13; Length 465;
Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2011 CCTACTAAAAATATATAAATTAGCTG 2036
Db 65 CCTACTAAAAATATATAAATTAGCTG 40
```

```
RESULT 16
US-10-027-632-295913/c
Sequence 295913, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 295913
LENGTH: 465
TYPE: DNA
ORGANISM: Human
US-10-027-632-295913
```

```
Query Match
Best Local Similarity 100.0%; Score 26; DB 13; Length 465;
Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAAAATATATAAATTAGCTG 2036

Db 65 CCTACTAAAAATATATAAATTAGCTG 40

RESULT 17  
US-10-027-632-66360/c

; Sequence 66360, Application US/10027632  
; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 66360

; LENGTH: 465

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-66360

Query Match 1.3%; Score 26; DB 15; Length 465;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAAAATATATAAATTAGCTG 2036

Db 65 CCTACTAAAAATATATAAATTAGCTG 40

RESULT 18  
US-10-027-632-67240/c

; Sequence 67240, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 67240

; LENGTH: 465

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-67240

Query Match 1.3%; Score 26; DB 15; Length 465;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAAAATATATAAATTAGCTG 2036

Db 65 CCTACTAAAAATATATAAATTAGCTG 40

RESULT 19  
US-10-027-632-295913/c

; Sequence 295913, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 295913

; LENGTH: 465

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-295913

Query Match 1.3%; Score 26; DB 15; Length 465;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2011 CCTACTAAAAATATATAAATTAGCTG 2036

Db 65 CCTACTAAAAATATATAAATTAGCTG 40

RESULT 20  
US-10-417-375-154/c

; Sequence 154, Application US/10417375

; Publication No. US20040219528A1

; GENERAL INFORMATION:

; APPLICANT: David W. Morris

; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer

; FILE REFERENCE: 529452001600

; CURRENT APPLICATION NUMBER: US/10/417,375

; CURRENT FILING DATE: 2003-04-15

; NUMBER OF SEQ ID NOS: 176

SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 154  
; LENGTH: 43329  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-417-375-154

Query Match 1.3%; Score 26; DB 18; Length 43329;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2011 CCTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 4419 CCTACTAAATATATAAATTAGCTG 4394

RESULT 21  
US-10-741-601-11354  
; Sequence 11354, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11354  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-11354

Query Match 1.2%; Score 25; DB 17; Length 201;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 73 CTACTAAATATATAAATTAGCTG 97

RESULT 22  
US-10-741-601-11355  
; Sequence 11355, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11355  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-11355

Query Match 1.2%; Score 25; DB 17; Length 201;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 10 CTACTAAATATATAAATTAGCTG 34

RESULT 23

US-10-741-601-14743/C  
; Sequence 14743, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14743  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-14743

Query Match 1.2%; Score 25; DB 17; Length 201;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 166 CTACTAAATATATAAATTAGCTG 142

RESULT 24  
US-09-764-891-6973  
; Sequence 6973, Application US/09764891  
; Publication No. US2003007808A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSEN et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6973  
; LENGTH: 213  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-6973

Query Match 1.2%; Score 25; DB 10; Length 213;  
Best Local Similarity 100.0%; Pred. No. 0.08;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 102 CTACTAAATATATAAATTAGCTG 126

RESULT 25  
US-10-242-535A-36118/C  
; Sequence 36118, Application US/10242535A  
; Publication No. US2004001363A1  
; GENERAL INFORMATION:  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; Prior application number: US 10/085,783  
; Prior filing date: 2002-02-28  
; Prior application number: US 60/305,340  
; Prior filing date: 2001-07-13  
; Prior application number: US 60/275,017  
; Prior filing date: 2001-03-12  
; Prior application number: US 60/271,955

```
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 36118
/ LENGTH: 230
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (17)..(17)
/ OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-36118
```

```
Query Match 1.2%; Score 25; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 226 CTACTAAAAATATATAAATTAGCTG 202
```

```
RESULT 26
US-10-085-783A-36118/c
/ Sequence 36118, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 36118
/ LENGTH: 230
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (17)..(17)
/ OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-36118
```

```
Query Match 1.2%; Score 25; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 226 CTACTAAAAATATATAAATTAGCTG 202
```

```
RESULT 27
US-10-242-535A-39800
/ Sequence 39800, Application US/10242535A
/ Publication No. US20040013663A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2005
/ CURRENT APPLICATION NUMBER: US/10/242,535A
/ PRIOR FILING DATE: 2002-09-12
/ PRIOR APPLICATION NUMBER: US 10/085,783
/ PRIOR FILING DATE: 2002-02-28
```

```
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 39800
/ LENGTH: 239
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (15)..(15)
/ OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-39800
```

```
Query Match 1.2%; Score 25; DB 16; Length 239;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 83 CTACTAAAAATATATAAATTAGCTG 107
```

```
RESULT 28
US-10-085-783A-39800
/ Sequence 39800, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 39800
/ LENGTH: 239
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (15)..(15)
/ OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-39800
```

```
Query Match 1.2%; Score 25; DB 16; Length 239;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 83 CTACTAAAAATATATAAATTAGCTG 107
```

```
RESULT 29
US-09-867-701-808
/ Sequence 808, Application US/09867701
/ Patent No. US20020132237A1
/ GENERAL INFORMATION:
/ APPLICANT: Aglate, Paul A.
/ APPLICANT: Jones, Robert
/ APPLICANT: Harlocker, Susan L.
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 808
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(282)
; OTHER INFORMATION: n = A,T,C or G
; US-09-867-701-808

Query Match      1.2%; Score 25; DB 9; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      99 CTACTAAAAATATATAAATTAGCTG 123

RESULT 30
US-10-040-739-673/C
; Sequence 673, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacoby, Kenneth
; McCoy, John
; Lavallee, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8824
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 673:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 673:

; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO OSTEOARTHRITIS
; TITLE OF INVENTION: AND METHODS RELATING TO OSTEOARTHRITIS
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9776
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Human
; US-10-242-535A-9776

Query Match      1.2%; Score 25; DB 13; Length 323;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      39 CTACTAAAAATATATAAATTAGCTG 15

RESULT 32
US-10-085-783A-9776/C
; Sequence 9776, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; Applicant: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9776
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Human
; US-10-085-783A-9776

Query Match      1.2%; Score 25; DB 16; Length 323;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      39 CTACTAAAAATATATAAATTAGCTG 15

RESULT 31
US-10-242-535A-9776/C
; Sequence 9776, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; Applicant: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9776
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Human
; US-10-242-535A-9776

Query Match      1.2%; Score 25; DB 13; Length 292;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      264 CTACTAAAAATATATAAATTAGCTG 240
```

Best Local Similarity 100.0%; Pred. No. 0.075;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||

Db 39 CTTACTAAAAATATATAAAATTAGCTG 15  
|||||

## RESULT 33

US-10-242-535A-14757  
; Sequence 14757, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 14757  
; LENGTH: 347  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-14757

Query Match 1.2%; Score 25; DB 16; Length 347;  
Best Local Similarity 100.0%; Pred. No. 0.074;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||

Db 180 CTTACTAAAAATATATAAAATTAGCTG 204  
|||||

## RESULT 34

US-10-085-783A-14757  
; Sequence 14757, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 14757  
; LENGTH: 347  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-14757

Query Match 1.2%; Score 25; DB 16; Length 347;  
Best Local Similarity 100.0%; Pred. No. 0.074;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||

Db 180 CTTACTAAAAATATATAAAATTAGCTG 204  
|||||

## RESULT 35

US-10-674-124A-23964  
; Sequence 23964, Application US/10674124A  
; Publication No. US2004019797A1  
; GENERAL INFORMATION:  
; APPLICANT: INOKO, Hidetoshi  
; APPLICANT: TAMAYA, Gen  
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE  
; FILE REFERENCE: ORIN-003CIP  
; CURRENT APPLICATION NUMBER: US/10/674,124A  
; PRIOR FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 10/257,511  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: PCT/JP00/07621  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: JP2000-112699  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: JP2002-327516  
; PRIOR FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: JP2002-383869  
; PRIOR FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 27110  
; SEQ ID NO 23964  
; LENGTH: 370  
; TYPE: DNA  
; ORGANISM: Homo sapiens

FEATURE: OTHER INFORMATION: Located on chromosome 19  
; FEATURE:  
; OTHER INFORMATION: Distance between a terminus base of telomere on  
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base  
; OTHER INFORMATION: sequence : 23198685  
; FEATURE:  
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of  
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and  
; OTHER INFORMATION: 5'-terminus of this base sequence : 156988  
US-10-674-124A-23964

Query Match 1.2%; Score 25; DB 18; Length 370;  
Best Local Similarity 100.0%; Pred. No. 0.073;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTTACTAAAAATATATAAAATTAGCTG 2036  
|||||

Db 55 CTTACTAAAAATATATAAAATTAGCTG 79  
|||||

## RESULT 36

US-10-027-632-19134/C  
; Sequence 19134, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218



```

:
:
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 19134
:
: LENGTH: 394
:
: TYPE: DNA
:
: ORGANISM: Human
:
: US-10-027-632-19134

```

```
Query Match      1.2%; Score 25; DB 13; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	2012	CTACTAA	AAATAT	AAAAAT	TAGCTG	2036
Db	155	CTACTAA	AAATAT	AAAAAT	TAGCTG	131

```

RESULT 37
US-10-027-632-19134/c
: Sequence 19134, Application US/10027632
: Publication No. US20030204075M8
GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: FILE REFERENCE: Polymorphisms in the Human Genome
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 19134
: LENGTH: 394
: TYPE: DNA
: ORGANISM: Human
: US-10-027-632-19134

```

Query Match	1.2%;	Score 25;	DB 15;	Length 394;
Best Local Similarity	100.0%;	Pred. NO. 0.073;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

RESULT 8  
US-09-795-668-1061/c  
; Sequence 1061, Application US/09795668  
; Patent No. US20020045577A1  
; GENERAL INFORMATION:  
; APPLICANT: Stefansson, Hreinn  
; APPLICANT: Steinhorsdottir, Valgerdur

```

:
: APPLICANT: Guiche, Jeffrey R.
: TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENES
: FILE REFERENCE: 2345, 2004-001
: CURRENT APPLICATION NUMBER: US/09/795,668
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: US 09/515,716
: PRIOR FILING DATE: 2000-02-28
: NUMBER OF SEQ ID NOS: 1531
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1061
:   LENGTH: 401
:   TYPE: DNA
: ORGANISM: Homo sapiens
:
: US-09-795-668-1061

```

Query Match	1.2%;	Score 25;	DB 9;	Length 401;
Best Local Similarity	100.0%;	Pred. No. 0.072;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 39
US-09-795-686-1061/c
: Sequence 1061, Application US/09795686
: Patent No. US20020094954A1
: GENERAL INFORMATION:
: APPLICANT: Stefansson, Hreinn
: APPLICANT: Stenlundorodottir, Valgerdur
: APPLICANT: Gulcher, Jeffrey R.
: TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
: FILE REFERENCE: 2345, 2005-001
: CURRENT APPLICATION NUMBER: US/09/795,686
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: US 09/515,715
: PRIOR FILING DATE: 2000-02-28
: NUMBER OF SEQ ID NOS: 1531
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1061
: LENGTH: 401
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-795-686-1061

```

Query Match	1.2%;	Score 25;	DB 9;	Length 401;
Best Local Similarity	100.0%;	Pred. No. 0.072;		
Matches 25; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	2012	CTACTAAATAATATAAAATTAGCTG	2036
Db	184	CTACTAAATAATATAAAATTAGCTG	160

```

RESULT 40
US-09-946-807-1061/c
: Sequence 1061, Application US/09946807
: Patent No. US20020165144A1
: GENERAL INFORMATION:
: APPLICANT: Stefansson, Heilm
: APPLICANT: Steinhilberdothlr, Valgerdur
: APPLICANT: Gulcher, Jeffrey R.
: TITLE OF INVENTION: HUMAN SCHIZOPHRENIA
: FILE REFERENCE: 2345..2004-001
: CURRENT APPLICATION NUMBER: US/09/946, 807
: CURRENT FILING DATE: 2001-09-05
: PRIOR APPLICATION NUMBER: US/09/795, 668
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: US 09/515, 716
: PRIOR FILING DATE: 2000-02-28
: NUMBER OF SEQ ID NOS: 1531
: SOFTWARE: FastQ for Windows Version 4.0

```



```

; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39808
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-39808

Query Match
Best Local Similarity 100.0%; Score 25; DB 16; Length 419;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 252 CTACTAAAAATATATAAATTAGCTG 276

RESULT 45
US-10-085-783A-39808
; Sequence 39808, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39808
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-39808

Query Match
Best Local Similarity 100.0%; Score 25; DB 16; Length 419;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 252 CTACTAAAAATATATAAATTAGCTG 276

RESULT 46
US-10-357-930-17263/C
; Sequence 17263, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
```

```

; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17263
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-17263

Query Match
Best Local Similarity 100.0%; Score 25; DB 18; Length 427;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 375 CTACTAAAAATATATAAATTAGCTG 351

RESULT 47
US-10-027-632-283895/C
; Sequence 283895, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283895
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-283895

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 430;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 233 CTACTAAAAATATATAAATTAGCTG 209
```

```
RESULT 48
US-10-027-632-283895/c
; Sequence 283895, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283895
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-283895

Query Match
Best Local Similarity 1.2%; Score 25; DB 15; Length 430;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTRACTAAAAATATAAAAATTAGCTG 2036
DB 233 CTRACTAAAAATATAAAAATTAGCTG 209

RESULT 49
US-09-918-995-34710
; Sequence 34710, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34710
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(434)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-34710

Query Match
Best Local Similarity 1.2%; Score 25; DB 10; Length 434;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTRACTAAAAATATAAAAATTAGCTG 2036
```

```
DB 171 CTRACTAAAAATATAAAAATTAGCTG 195

RESULT 50
US-10-674-124A-15261/c
; Sequence 15261, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 15261
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Located on chromosome 9
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence : 79581780
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 94980
US-10-674-124A-15261

Query Match
Best Local Similarity 1.2%; Score 25; DB 18; Length 450;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTRACTAAAAATATAAAAATTAGCTG 2036
DB 268 CTRACTAAAAATATAAAAATTAGCTG 244

RESULT 51
US-10-027-632-270729/c
; Sequence 270729, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

```

; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 270729
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-270729

Query Match
Best Local Similarity 1.2%; Score 25; DB 13; Length 461;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 236 CTAATAAATAATTAATTAAGCTG 212

RESULT 52
US-10-027-632-270729/c
; Sequence 270729, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 270729
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-270729

Query Match
Best Local Similarity 1.2%; Score 25; DB 15; Length 461;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 236 CTAATAAATAATTAATTAAGCTG 212

RESULT 53
US-10-674-124A-15436
; Sequence 15436, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidecoshi
; APPLICANT: TAMAYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
```

```

; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 15436
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chr9.fa.07fz.111287695
; FEATURE:
; OTHER INFORMATION: Located on chromosome 9
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 152146
US-10-674-124A-15436

Query Match
Best Local Similarity 1.2%; Score 25; DB 18; Length 461;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAATTAAGCTG 2036
DB 205 CTAATAAATAATTAATTAAGCTG 229

RESULT 54
US-10-674-124A-1319/c
; Sequence 1319, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidecoshi
; APPLICANT: TAMAYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 1319
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: ALI21987.3_66483
; FEATURE:
; OTHER INFORMATION: Located on chromosome 1
```

```
/ FEATURE:
/ OTHER INFORMATION: Distance between a terminus base of telomere on
/ OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
/ OTHER INFORMATION: sequence : 160932243
/ FEATURE:
/ OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
/ OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
/ OTHER INFORMATION: 5'-terminus of this base sequence : 8635
US-10-674-124A-1319

Query Match
Best Local Similarity 1.2%; Score 25; DB 18; Length 463;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 409 CTAATAAATATATAAATTAGCTG 385

RESULT 55
US-10-357-930-47086/C
/ Sequence 47086, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Edege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-0078CN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 47086
/ LENGTH: 467
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-357-930-47086

Query Match
Best Local Similarity 1.2%; Score 25; DB 18; Length 467;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 402 CTAATAAATATATAAATTAGCTG 378

RESULT 56
US-10-674-124A-4488/C
/ Sequence 4488, Application US/10674124A
/ Publication No. US20040197797A1
/ GENERAL INFORMATION:
/ APPLICANT: INOKO, Hidetoshi
/ APPLICANT: TAMITA, Gen
/ TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
/ TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
```

```
/ FILE REFERENCE: ORIN-003CIP
/ CURRENT APPLICATION NUMBER: US/10/674,124A
/ CURRENT FILING DATE: 2003-09-26
/ PRIOR APPLICATION NUMBER: 10/257,511
/ PRIOR FILING DATE: 2003-03-07
/ PRIOR APPLICATION NUMBER: PCT/JP00/07621
/ PRIOR FILING DATE: 2000-10-30
/ PRIOR APPLICATION NUMBER: JP2000-112699
/ PRIOR FILING DATE: 2000-04-13
/ PRIOR APPLICATION NUMBER: JP2002-327516
/ PRIOR FILING DATE: 2002-09-28
/ PRIOR APPLICATION NUMBER: JP2002-383869
/ PRIOR FILING DATE: 2002-12-09
/ NUMBER OF SEQ ID NOS: 27110
/ SEQ ID NO 4488
/ LENGTH: 470
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: chr2.fa.07fz.245257698
/ FEATURE:
/ OTHER INFORMATION: Located on chromosome 2
/ FEATURE:
/ OTHER INFORMATION: Distance between a terminus base of telomere on
/ OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
/ OTHER INFORMATION: sequence : 233569065
/ FEATURE:
/ OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
/ OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
/ OTHER INFORMATION: 5'-terminus of this base sequence : 143819
US-10-674-124A-4488

Query Match
Best Local Similarity 1.2%; Score 25; DB 18; Length 470;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 336 CTAATAAATATATAAATTAGCTG 312

RESULT 57
US-09-867-701-6710/C
/ Sequence 6710, Application US/09867701
/ Patent No. US2002013237A1
/ GENERAL INFORMATION:
/ APPLICANT: Aglate, Paul A.
/ APPLICANT: Jones, Robert
/ APPLICANT: Harlocker, Susan L.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.497
/ CURRENT APPLICATION NUMBER: US/09/867,701
/ CURRENT FILING DATE: 2001-05-29
/ NUMBER OF SEQ ID NOS: 10912
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6710
/ LENGTH: 472
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-867-701-6710

Query Match
Best Local Similarity 1.2%; Score 25; DB 9; Length 472;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 171 CTAATAAATATATAAATTAGCTG 147

RESULT 58
US-09-918-995-2669/C
```

```
; Sequence 2669, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2669
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(474)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-2669

Query Match          1.2%; Score 25; DB 10; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTAATAAATATATAAATTAGCTG 2036
DB      104 CTAATAAATATATAAATTAGCTG 80

RESULT 59
US-10-242-535A-24474
; Sequence 24474, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24474
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Human
; US-10-242-535A-24474

Query Match          1.2%; Score 25; DB 16; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTAATAAATATATAAATTAGCTG 2036
DB      401 CTAATAAATATATAAATTAGCTG 425

RESULT 60
US-10-085-783A-24474
; Sequence 24474, Application US/10085783A
; Publication No. US2004003781A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24474
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Human
; US-10-085-783A-24474

Query Match          1.2%; Score 25; DB 16; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTAATAAATATATAAATTAGCTG 2036
DB      401 CTAATAAATATATAAATTAGCTG 425

RESULT 61
US-09-918-995-15293
; Sequence 15293, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15293
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(486)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-15293

Query Match          1.2%; Score 25; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTAATAAATATATAAATTAGCTG 2036
DB      309 CTAATAAATATATAAATTAGCTG 333

RESULT 62
US-09-918-995-34441
; Sequence 34441, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
```

```

; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34441
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(488)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-34441

Query Match
Best Local Similarity 1.2%; Score 25; DB 10; Length 488;
Pred. No. 0.07; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

Qy 2012 CTAATAAATATAAATTAGCTG 2036
Db 232 CTAATAAATATAAATTAGCTG 256

RESULT 63
US-10-027-632-43038/c
; Sequence 43038, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43038
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43038

Query Match
Best Local Similarity 1.2%; Score 25; DB 13; Length 497;
Pred. No. 0.07; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

Qy 2012 CTAATAAATATAAATTAGCTG 2036
Db 248 CTAATAAATATAAATTAGCTG 224

RESULT 64
US-10-027-632-72533/c
; Sequence 72533, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```

; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43038
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43038

Query Match
Best Local Similarity 1.2%; Score 25; DB 13; Length 497;
Pred. No. 0.07; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

Qy 2012 CTAATAAATATAAATTAGCTG 2036
Db 248 CTAATAAATATAAATTAGCTG 224

RESULT 65
US-10-027-632-43038/c
; Sequence 43038, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43038
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43038

Query Match
Best Local Similarity 1.2%; Score 25; DB 15; Length 497;
Pred. No. 0.07; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

Qy 2012 CTAATAAATATAAATTAGCTG 2036
```



Db 248 CTAAGTAAATATTAATTTAGCTG 224

RESULT 66  
US-10-027-632-72533/c  
; Sequence 72533, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 72533  
; LENGTH: 497  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-72533

Query Match 1.2%; Score 25; DB 15; Length 497;  
Best Local Similarity 100.0%; Pred. No. 0.07;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAAGTAAATATTAATTTAGCTG 2036  
Db 248 CTAAGTAAATATTAATTTAGCTG 224

RESULT 67  
US-10-240-425-252/c  
; Sequence 252, Application US/10240425  
; Publication No. US20040033502A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Amanda  
; APPLICANT: Boland, Joseph F.  
; APPLICANT: Lord, Reginald V.  
; APPLICANT: Alvarez, Chris  
; APPLICANT: Metzel, Jon C.  
; APPLICANT: Scheff, Uwe  
; APPLICANT: Vockley, Joseph G.  
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue  
; FILE REFERENCE: 44921-5026  
; CURRENT APPLICATION NUMBER: US/10/240,425  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: PCT/US01/09847  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 60/193,446  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 1588  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 252  
; LENGTH: 503  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AA767440  
US-10-240-425-252

Query Match 1.2%; Score 25; DB 16; Length 503;  
Best Local Similarity 100.0%; Pred. No. 0.07;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAAGTAAATATTAATTTAGCTG 2036  
Db 183 CTAAGTAAATATTAATTTAGCTG 159

RESULT 68  
US-10-027-632-265243/c  
; Sequence 265243, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 265243  
; LENGTH: 508  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-265243

Query Match 1.2%; Score 25; DB 13; Length 508;  
Best Local Similarity 100.0%; Pred. No. 0.07;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTAAGTAAATATTAATTTAGCTG 2036  
Db 94 CTAAGTAAATATTAATTTAGCTG 70

RESULT 69  
US-10-027-632-265243/c  
; Sequence 265243, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24

```
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 265243
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-265243

Query Match
Best Local Similarity 100.0%; Score 25; DB 15; Length 508;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      94 CTACTAAAAATATATAAATTAGCTG 70

RESULT 70
US-10-102-524-1199
; Sequence 1199, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1199
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 504
; OTHER INFORMATION: n = A,T,C or G
US-10-102-524-1199

Query Match
Best Local Similarity 100.0%; Score 25; DB 15; Length 515;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      23 CTACTAAAAATATATAAATTAGCTG 47

RESULT 71
US-10-027-632-289987
; Sequence 289987, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289987
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-289987

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 526;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      242 CTACTAAAAATATATAAATTAGCTG 266

RESULT 72
US-10-027-632-289987
; Sequence 289987, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289987
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-289987

Query Match
Best Local Similarity 100.0%; Score 25; DB 15; Length 526;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATATATAAATTAGCTG 2036
Db      242 CTACTAAAAATATATAAATTAGCTG 266

RESULT 73
US-10-027-632-270728/c
; Sequence 270728, Application US/10027632
```

```
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 270728
LENGTH: 536
TYPE: DNA
ORGANISM: Human
US-10-027-632-270728

Query Match
Best Local Similarity 1.2%; Score 25; DB 13; Length 536;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 314 CTACTAAAAATATATAAATTAGCTG 290

RESULT 74
US-10-027-632-270728/c
Sequence 270728, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 270728
LENGTH: 536
TYPE: DNA
ORGANISM: Human
US-10-027-632-270728

Query Match
1.2%; Score 25; DB 15; Length 536;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 314 CTACTAAAAATATATAAATTAGCTG 290

RESULT 75
US-10-027-632-242121
Sequence 242121, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 242121
LENGTH: 546
TYPE: DNA
ORGANISM: Human
US-10-027-632-242121

Query Match
1.2%; Score 25; DB 13; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036
Db 59 CTACTAAAAATATATAAATTAGCTG 83

RESULT 76
US-10-027-632-242121
Sequence 242121, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
```

```
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242121
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-242121
```

```
Query Match          1.2%; Score 25; DB 15; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2012 CTAATAAATATTAATAGCTG 2036
Db      59  CTAATAAATATTAATAGCTG 83
```

RESULT 77

```
US-10-027-632-86815
; Sequence 86815, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

```
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86815
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86815
```

```
Query Match          1.2%; Score 25; DB 13; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2012 CTAATAAATATTAATAGCTG 2036
Db      295 CTAATAAATATTAATAGCTG 319
```

RESULT 78

```
US-10-027-632-86816
; Sequence 86816, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86816
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86816
```

```
Query Match          1.2%; Score 25; DB 13; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2012 CTAATAAATATTAATAGCTG 2036
Db      295 CTAATAAATATTAATAGCTG 319
```

RESULT 79

```
US-10-027-632-316485
; Sequence 316485, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

```
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316485
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-316485
```

```
Query Match          1.2%; Score 25; DB 13; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2012 CTAATAAATATTAATAGCTG 2036
Db      295 CTAATAAATATTAATAGCTG 319
```

RESULT 80

```
US-10-027-632-316485
; Sequence 316485, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
US-10-027-632-316486
; Sequence 316486, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316486
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-316486

Query Match          1.2%; Score 25; DB 13; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATAATTAATAATTAGCTG 2036
DB 295 CTACTAAATAATTAATAATTAGCTG 319

RESULT 81
US-10-027-632-86815
; Sequence 86815, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86815
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86815
```

```
Query Match          1.2%; Score 25; DB 15; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATAATTAATAATTAGCTG 2036
DB 295 CTACTAAATAATTAATAATTAGCTG 319

RESULT 82
US-10-027-632-86816
; Sequence 86816, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86816
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86816

Query Match          1.2%; Score 25; DB 15; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATAATTAATAATTAGCTG 2036
DB 295 CTACTAAATAATTAATAATTAGCTG 319

RESULT 83
US-10-027-632-316485
; Sequence 316485, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
```

```
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316485
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-316485

Query Match          1.2%; Score 25; DB 15; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 295 CTACTAAAAATATATAAATTAGCTG 319

RESULT 84
US-10-027-632-316486
; Sequence 316486, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316486
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-316486

Query Match          1.2%; Score 25; DB 15; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 295 CTACTAAAAATATATAAATTAGCTG 319

RESULT 85
US-10-027-632-283621
; Sequence 283621, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283621
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-283621

Query Match          1.2%; Score 25; DB 15; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283621
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-283621

Query Match          1.2%; Score 25; DB 13; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 141 CTACTAAAAATATATAAATTAGCTG 165

RESULT 86
US-10-027-632-283621
; Sequence 283621, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283621
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-283621

Query Match          1.2%; Score 25; DB 15; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036
DB 141 CTACTAAAAATATATAAATTAGCTG 165
```

```
RESULT 87
US-10-027-632-219836/c
; Sequence 219836, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219836
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-219836

Query Match          1.2%; Score 25; DB 13; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATTAATAATTAGCTG 2036
DB      374 CTACTAAAAATTAATAATTAGCTG 350

RESULT 88
US-10-027-632-219836/c
; Sequence 219836, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219836
; LENGTH: 559
; TYPE: DNA
```

```
; ORGANISM: Human
US-10-027-632-219836

Query Match          1.2%; Score 25; DB 15; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATTAATAATTAGCTG 2036
DB      374 CTACTAAAAATTAATAATTAGCTG 350

RESULT 89
US-10-027-632-268734
; Sequence 268734, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268734
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-268734

Query Match          1.2%; Score 25; DB 13; Length 561;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAAAATTAATAATTAGCTG 2036
DB      17 CTACTAAAAATTAATAATTAGCTG 41

RESULT 90
US-10-027-632-268735
; Sequence 268735, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

```
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 268735
/ LENGTH: 561
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-268735

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 561;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTAATAAAATATATAAATTAGCTG 2036
17 CTAATAAAATATATAAATTAGCTG 41

RESULT 91
US-10-027-632-268734
/ Sequence 268734, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 268734
/ LENGTH: 561
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-268734

Query Match
Best Local Similarity 100.0%; Score 25; DB 15; Length 561;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTAATAAAATATATAAATTAGCTG 2036
17 CTAATAAAATATATAAATTAGCTG 41

RESULT 92
US-10-027-632-268735
/ Sequence 268735, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 268735
/ LENGTH: 561
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-268735

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 561;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTAATAAAATATATAAATTAGCTG 2036
17 CTAATAAAATATATAAATTAGCTG 41

RESULT 93
US-10-027-632-105456
/ Sequence 105456, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 105456
/ LENGTH: 564
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-105456

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 564;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTAATAAAATATATAAATTAGCTG 2036
17 CTAATAAAATATATAAATTAGCTG 41
```

```
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 268735
/ LENGTH: 561
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-268735

Query Match
Best Local Similarity 100.0%; Score 25; DB 15; Length 561;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTAATAAAATATATAAATTAGCTG 2036
17 CTAATAAAATATATAAATTAGCTG 41

RESULT 93
US-10-027-632-105456
/ Sequence 105456, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 105456
/ LENGTH: 564
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-105456

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 564;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2012 CTAATAAAATATATAAATTAGCTG 2036
17 CTAATAAAATATATAAATTAGCTG 41
```



Db 225 CTACTAAAAATATAAAATTAGCTG 249

RESULT 94  
US-10-027-632-105456  
; Sequence 105456, Application US/10027632  
; Publication No. US20030204075A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 105456

LENGTH: 564

TYPE: DNA

ORGANISM: Human

US-10-027-632-105456

Query Match 1.2% Score 25; DB 15; Length 564;  
Best Local Similarity 100.0%; Pred. No. 0.069;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATAAAATTAGCTG 2036

Db 225 CTACTAAAAATATAAAATTAGCTG 249

RESULT 95  
US-10-027-632-257015/c  
; Sequence 257015, Application US/10027632  
; Publication No. US20020198371A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 257015

LENGTH: 573

TYPE: DNA

ORGANISM: Human

US-10-027-632-257015

Query Match 1.2% Score 25; DB 13; Length 573;  
Best Local Similarity 100.0%; Pred. No. 0.069;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATAAAATTAGCTG 2036

Db 380 CTACTAAAAATATAAAATTAGCTG 356

RESULT 96  
US-10-027-632-257015/c  
; Sequence 257015, Application US/10027632  
; Publication No. US20030204075A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 257015

LENGTH: 573

TYPE: DNA

ORGANISM: Human

US-10-027-632-257015

Query Match 1.2% Score 25; DB 15; Length 573;  
Best Local Similarity 100.0%; Pred. No. 0.069;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATAAAATTAGCTG 2036

Db 380 CTACTAAAAATATAAAATTAGCTG 356

RESULT 97  
US-10-027-632-78458  
; Sequence 78458, Application US/10027632  
; Publication No. US20020198371A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

```

; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78458
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(574)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-78458
```

```
Query Match
Best Local Similarity 1.2%; Score 25; DB 13; Length 574;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 193 CTAATAAATATATAAATTAGCTG 217
```

```

RESULT 98
US-10-027-632-78458
; Sequence 78458, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78458
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(574)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-78458
```

```
Query Match
Best Local Similarity 1.2%; Score 25; DB 15; Length 574;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 193 CTAATAAATATATAAATTAGCTG 217
```

```

RESULT 99
US-10-027-632-60104
; Sequence 60104, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60104
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-60104
```

```
Query Match
Best Local Similarity 1.2%; Score 25; DB 13; Length 575;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2012 CTAATAAATATATAAATTAGCTG 2036
Db 301 CTAATAAATATATAAATTAGCTG 325
```

```

RESULT 100
US-10-027-632-60589
; Sequence 60589, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60589
; LENGTH: 575
```

TYPE: DNA  
ORGANISM: Human  
US-10-027-632-60589

Query Match 1.2%; Score 25; DB 13; Length 575;  
Best Local Similarity 100.0%; Pred. No. 0.068;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATAATTAAGCTG 2036  
|||||  
DB 301 CTAATAAATAATTAAGCTG 325

Search completed: January 5, 2005, 01:12:11  
UOB time : 1095 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 4, 2005, 18:04:11 ; Search time 5765 Seconds  
(without alignment)

12869.262 Million cell updates/sec

Title: US-09-508-658B-1

Perfect score: 2036

Sequence: 1 agaccggggagagcgggcg9.....aaaataaaaaaattagctg 2036

Scoring table: OLIGO=NUC  
Gapop 60.0, Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 0

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 200 summaries

Database :  
EST:  
1: gb\_ests1.\*  
2: gb\_ests2.\*  
3: gb\_hic.\*  
4: gb\_ests3.\*  
5: gb\_ests4.\*  
6: gb\_ests5.\*  
7: gb\_ests6.\*  
8: gb\_gse81.\*  
9: gb\_gse82.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	747	36.7	1506	9	AY419550 Homo sapi
2	454	22.3	642	5	BX112453 BX112453
3	229	11.2	1474	9	AY419551 Pan trogl
4	157	7.7	403	1	AA742555 nx30c11.s
5	69	3.4	200	9	HSDML04M6
6	65	3.2	350	8	AQ474914
7	32	1.6	451	2	BE627856
8	32	1.6	1512	9	AY419552 Mus muscu
9	30	1.5	384	1	AI552580 vx26b07.x
10	30	1.5	420	9	CG513175 OST66421
11	30	1.5	511	2	BE630816 uu49c07.x
12	26	1.3	459	1	AA703834
13	26	1.3	478	8	BZ277407
14	26	1.3	479	8	AQ239234
15	26	1.3	512	9	CG585187
16	26	1.3	634	6	BY745644
17	26	1.3	662	4	BI667421
18	26	1.3	680	9	AG185660
19	26	1.3	715	9	AG185665
20	26	1.3	742	7	CN315153
21	26	1.3	923	8	BQ431008
22	25	1.2	119	5	AO530389
23	25	1.2	138	8	AO541617
24	25	1.2	149	2	BF843049

25	25	1.2	150	7	T70055
26	25	1.2	152	2	AM997297
27	25	1.2	162	7	T70121
28	25	1.2	179	2	BF915807
29	25	1.2	184	1	AA664564
30	25	1.2	187	9	AG204713
31	25	1.2	190	2	BF807993
32	25	1.2	194	2	BF883245
33	25	1.2	200	1	AI193409
34	25	1.2	211	1	AA626595
35	25	1.2	232	4	BI007053
36	25	1.2	235	4	BG615666
37	25	1.2	236	2	AM854958
38	25	1.2	237	4	BS531549
39	25	1.2	248	1	AA847405
40	25	1.2	252	8	BS2975
41	25	1.2	255	2	BF851782
42	25	1.2	256	1	AA618344
43	25	1.2	256	7	P26713
44	25	1.2	256	7	N73518
45	25	1.2	260	6	CD516253
46	25	1.2	266	1	AA847504
47	25	1.2	270	1	AA829490
48	25	1.2	272	1	AA468966
49	25	1.2	279	6	CD357167
50	25	1.2	279	7	T59425
51	25	1.2	280	2	BE043767
52	25	1.2	282	7	T59481
53	25	1.2	285	7	R82328
54	25	1.2	287	4	BI062031
55	25	1.2	292	2	AM770166
56	25	1.2	293	4	BI031941
57	25	1.2	294	2	AM832979
58	25	1.2	302	8	AA719115
59	25	1.2	303	8	BF4377
60	25	1.2	307	8	AQ585018
61	25	1.2	308	8	AO104891
62	25	1.2	310	8	AO083275
63	25	1.2	312	8	BI109463
64	25	1.2	312	8	AQ029448
65	25	1.2	317	1	AA625620
66	25	1.2	317	2	BF836228
67	25	1.2	319	1	AA632845
68	25	1.2	320	1	BE008292
69	25	1.2	326	7	R00494
70	25	1.2	329	2	AM938545
71	25	1.2	331	7	R08502
72	25	1.2	338	2	BF081268
73	25	1.2	339	1	AV733997
74	25	1.2	339	5	BX480803
75	25	1.2	339	8	BS3941
76	25	1.2	346	2	BE071881
77	25	1.2	346	8	B90305
78	25	1.2	347	5	BX953655
79	25	1.2	348	1	AA837035
80	25	1.2	348	2	BF081260
81	25	1.2	349	7	R6389
82	25	1.2	352	7	HS8252
83	25	1.2	353	1	AA225629
84	25	1.2	354	8	AO705989
85	25	1.2	357	2	BF820331
86	25	1.2	357	8	AO374407
87	25	1.2	357	8	AO507012
88	25	1.2	359	2	BE185887
89	25	1.2	361	8	BE205649
90	25	1.2	363	1	AI124635
91	25	1.2	366	1	AI283938
92	25	1.2	368	4	BM671780
93	25	1.2	369	7	R08675
94	25	1.2	371	1	AA646212
95	25	1.2	372	1	AA195886
96	25	1.2	374	2	BE090515
97	25	1.2	376	1	AA640617

T70055	yc17c08.t1
AM997297	RC2-BN004
T70121	yc17c08.t1
BF915807	IL3-UT011
AA664564	nu71d07.s
AG204713	Pan trogl
BF807993	RC3-C1004
BF883245	QV3-ET017
AI193409	qes8c03.x
AA626595	ab54g03.x
BI007053	RC5-RT005
BG615666	602643025
AM854958	PM0-CT026
BS531549	602559760
AA847405	ce18f09.s
BS2975	CIT-HSP-200
BF851782	CM3-EN007
AA618344	ng18a03.s
P26713	HSPD14290.H
N73518	z49e03.t1
CD516253	AGENCOURT
AA847504	ce19b03.s
AA829490	od06g04.s
AA468966	nc79b09.x
CD357167	AGENCOURT
T59425	yb63f04.t1
BE043767	hk43d02.y
T59481	yb63f04.t1
R82328	y118a03.t1
BI062031	IL3-UT011
AM770166	hk59b06.x
BI031941	IL5-MT026
AM832979	RC3-TT000
AA719115	zh33f01.s
BF4377	CIT-HSP-202
AQ585018	RPCI-11-4
AO104891	HS 2166.A
AO083275	RPCI11-57
BI109463	BX109463
AQ029448	RPCI11-38
AA625620	ad10a11.s
BF836228	QV3-HT101
AA632845	np84f02.s
BE008292	CM2-GN022
R00494	ye74b03.t1
AM938545	PM4-DT006
R08502	ye95f08.t1
BF081268	RC6-AN006
AV733997	AV733997
BX480803	DKF2P666E
BS3941	RPCI11-21E2
BE071881	RC2-BT052
B90305	CIT-HSP-217
BX953655	DKF2P7810
AA837035	CG09E11.s
BF081260	RC6-AN006
R6389	y131e03.t1
HS8252	y16b04.t1
AA225629	nc08a04.x
AQ705989	HS 5556.A
BF820331	CM0-RT001
AO374407	RPCI11-15
AQ507012	RPCI-11-3
BE185887	IL5-HT03
BZ205649	WHA6G09R
AI124635	am59g02.x
AI283938	qt71e11.x
BM671780	UI-E-COI-
R08675	ye20g08.t1
AA646212	AV646212
AA195886	zp92g10.x
BE090515	RC6-BT071
AA640617	nr22a01.x

```
98 25 1.2 376 2 AM86827
C 99 25 1.2 378 1 AO106283
C 100 25 1.2 379 1 AA58686
C 101 25 1.2 380 8 AQ098567
C 102 25 1.2 382 6 CD519195
C 103 25 1.2 383 2 BF854196
C 104 25 1.2 384 8 AO526521
C 105 25 1.2 384 1 AA551313
C 106 25 1.2 384 4 BG897176
C 107 25 1.2 385 8 BS6864
C 108 25 1.2 386 1 AI039539
C 109 25 1.2 386 5 BU589955
C 110 25 1.2 387 2 BF870739
C 111 25 1.2 388 4 BG990105
C 112 25 1.2 389 2 BE827710
C 113 25 1.2 390 2 BF771417
C 114 25 1.2 390 8 AQ074473
C 115 25 1.2 391 1 AA932512
C 116 25 1.2 393 2 AM264901
C 117 25 1.2 393 4 BG007276
C 118 25 1.2 394 1 AA397472
C 119 25 1.2 396 1 AI470091
C 120 25 1.2 396 1 AA487219
C 121 25 1.2 403 6 CB178809
C 122 25 1.2 404 8 AQ054322
C 123 25 1.2 405 1 AI355873
C 124 25 1.2 405 7 T99082
C 125 25 1.2 406 1 AA531270
C 126 25 1.2 406 4 BM828564
C 127 25 1.2 408 1 AA226382
C 128 25 1.2 408 2 BF837053
C 129 25 1.2 409 7 H68920
C 130 25 1.2 411 2 AM754471
C 131 25 1.2 411 8 AQ115207
C 132 25 1.2 412 1 AA769559
C 133 25 1.2 412 1 AI167842
C 134 25 1.2 412 1 AA487079
C 135 25 1.2 412 5 BX484919
C 136 25 1.2 412 8 AQ060163
C 137 25 1.2 414 7 CN258820
C 138 25 1.2 414 8 AQ202181
C 139 25 1.2 415 8 B95419
C 140 25 1.2 416 1 AA847585
C 141 25 1.2 417 8 AQ198862
C 142 25 1.2 418 1 AA584148
C 143 25 1.2 419 1 AA578907
C 144 25 1.2 419 2 AM170670
C 145 25 1.2 420 1 AI284108
C 146 25 1.2 420 1 AI732096
C 147 25 1.2 421 1 AI812082
C 148 25 1.2 422 8 AO526786
C 149 25 1.2 424 1 AA468322
C 150 25 1.2 424 8 AQ154163
C 151 25 1.2 425 8 B32619
C 152 25 1.2 426 1 AI357769
C 153 25 1.2 427 1 AI732139
C 154 25 1.2 428 8 AQ184503
C 155 25 1.2 429 2 AM802856
C 156 25 1.2 429 4 BG007563
C 157 25 1.2 430 1 AI471595
C 158 25 1.2 430 1 AI476313
C 159 25 1.2 430 8 AO056563
C 160 25 1.2 430 8 AO208536
C 161 25 1.2 431 8 AQ176936
C 162 25 1.2 433 8 AQ412573
C 163 25 1.2 435 2 AM832972
C 164 25 1.2 437 7 T59276
C 165 25 1.2 437 7 H10143
C 166 25 1.2 439 1 AA206858
C 167 25 1.2 440 1 AA586581
C 168 25 1.2 440 5 BX492504
C 169 25 1.2 440 8 AQ062372
C 170 25 1.2 441 1 AJ574548
```

```
AM86827 RCT-OT008
AO106283 HS_3051_B
AA58686 HG7005_-8
AQ098567 HS_3050_A
CD519195 AGENCOURT
BF854196 NMR2-EN009
AO526521 HS_5309_B
AA551313 HK7402_-8
BG897176 HOA60-1-P
BS6864 CIT-HSP-201
AI039539 OX37C06_-8
BU589955 AGENCOURT
BF870739 IL3-ET011
BG990105 RCO-HT009
BE827710 RCT-ET001
BF771417 IL5-IT002
AQ074473 CIT-HSP-2
AA932512 OOG1C05_-8
AA264901 X06A06_-X
BG007276 IL5-GN023
AA397472 HG65H10_-1
AI470091 LJ90603_-X
AA487219 AB19911_-8
CB178809 I440F09_-X
AQ054322 CIT-HSP-2
AO54322 CIT-HSP-2
T99082 Y6F7603_81
AA531270 NJ09C08_-8
BM828564 K-EST0101
AA226382 NC20604_-8
BF837053 RC6-HT101
H68920 YR85C05_81
AM754471 OYA-CT036
AQ115207 RPTC11-54
AA769559 n239C06_-8
AI167842 O290907_-X
AA487079 AB19912_-X
BX484919 DKF2P686B
AQ060163 CITBI-EI-
CN258820 170006000
AQ202181 RPTC11-60
B95419 CIT-HSP-217
AA847585 Oe49P10_-8
AQ198862 RPTC11-59
AA584148 n005P01_-8
AA578907 n26603_-8
AM170670 X185G08_-X
AI284108 Q72604_-X
AI732096 NC20604_-X
AI812082 LW77A09_-X
AO526786 HS_5218_A
AA468322 NC77A08_-8
AQ154163 HS_2169_A
B32619 HS-1015-B1-
B32619 HS-1015-B1-
AI357769 GU98D09_-X
AI732139 nC08A04_-X
AQ184503 HS_2198_B
AM802856 IL2-UM007
BG007563 RCT-GN026
AI471595 tA15G07_-X
AI476313 tA15C07_-X
AO056563 CIT-HSP-2
AO208536 HS_3237_A
AQ176936 HS_3208_A
AQ412573 RCTI-11-1
AM832972 RCT-TT000
T59276 YC02B09_-X1
H10143 YMO5A11_81
AA206858 ZG80F07_8
AA586581 n056C05_-8
BX492504 DKF2P781G
AQ062372 CIT-HSP-2
AJ574548 AJ574548
```

```
C 171 25 1.2 444 1 AA634209
C 172 25 1.2 444 2 BF882222
C 173 25 1.2 444 2 BF882223
C 174 25 1.2 444 2 AM514615
C 175 25 1.2 447 1 AU159116
C 176 25 1.2 448 6 CA943866
C 177 25 1.2 448 7 H60307
C 178 25 1.2 448 8 AO117690
C 179 25 1.2 449 2 AM833031
C 180 25 1.2 449 8 AQ12765
C 181 25 1.2 450 1 AI014920
C 182 25 1.2 450 2 BE618924
C 183 25 1.2 451 1 AA156541
C 184 25 1.2 451 8 AQ262949
C 185 25 1.2 452 5 BX486739
C 186 25 1.2 453 1 AI568009
C 187 25 1.2 453 8 AQ242362
C 188 25 1.2 454 5 BM998336
C 189 25 1.2 456 1 AI493665
C 190 25 1.2 456 1 AI890385
C 191 25 1.2 456 2 BF856040
C 192 25 1.2 456 6 CD236938
C 193 25 1.2 456 8 AQ784022
C 194 25 1.2 456 8 BZ609705
C 195 25 1.2 457 7 H54640
C 196 25 1.2 459 1 AI678754
C 197 25 1.2 459 7 H58595
C 198 25 1.2 463 8 AQ623886
C 199 25 1.2 464 1 AI734070
C 200 25 1.2 464 1 AI820951
```

## ALIGNMENTS

```
RESULT 1
AY419550
LOCUS
DEFINITION
AY419550
ACCESSION
AY419550
VERSION
AY419550.1 GI:39775507
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1506)
Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Interfing nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
AUTHORS
2 (bases 1 to 1506)
Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
```

```
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
```

## FEATURES

```
source
1..1506
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1506
/gene="AIRE"
```

gene

ORIGIN /locus\_tag="HCM6928"

Query Match 36.7%; Score 747; DB 9; Length 1506;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 747; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

269 GAGAGCCTTCATCTTGAAGAAAAAGAGGGCTGCCCCCAGGCTTCCAGCCCTCTGTGCTC 328  
DB 1 GAGAGCCTTCATCTTGAAGAAAAAGAGGGCTGCCCCCAGGCTTCCAGCCCTCTGTGCTC 60  
329 TGGAGTCTGACCCAGGACTCCAGCCATCTTGAAGCTTCTGAGAGGGTGTGCTTCAAGAGC 388  
DB 61 TGGAGTCTGACCCAGGACTCCAGCCATCTTGAAGCTTCTGAGAGGGTGTGCTTCAAGAGC 120  
389 TACAACCTGAGAGGCTATGAGCCGCTGAGCCCATCTGAGACAGCTTCCCAAGATGTG 448  
DB 121 TACAACCTGAGAGGCTATGAGCCGCTGAGCCCATCTGAGACAGCTTCCCAAGATGTG 180  
449 GACCTTCAGCAGCCCGGAG 508  
DB 181 GACCTTCAGCAGCCCGGAG 240  
509 CCACCCAGACTCCCAAG 568  
DB 241 CCACCCAGACTCCCAAG 300  
569 GCCCTGACTCCAAAGGGGCAACCGGAGCCAGAGCTTCAATGAGAGCCAAAGCCCGCCAG 628  
DB 301 GCCCTGACTCCAAAGGGGCAACCGGAGCCAGAGCTTCAATGAGAGCCAAAGCCCGCCAG 360  
629 AAGCCGGAAG 688  
DB 361 AAGCCGGAAG 420  
689 TCAGCTTCAGTCCAG 748  
DB 421 TCAGCTTCAGTCCAG 480  
749 GCCGTGAG 808  
DB 481 GCCGTGAG 540  
809 CAGGTTGGGGGAG 868  
DB 541 CAGGTTGGGGGAG 600  
869 AAGGCCCCGAG 928  
DB 601 AAGGCCCCGAG 660  
929 GCCCCGAGTGAAG 988  
DB 661 GCCCCGAGTGAAG 720  
989 CTCGCCAGTGAAG 1015  
DB 721 CTCGCCAGTGAAG 747

RESULT 2  
BX112453 642 bp mRNA linear EST 07-FEB-2003  
LOCUS BX112453 NC1 CGAP GC4 Homo sapiens cDNA clone IMAGE998231165;  
DEFINITION IMAGE:1257620, mRNA sequence.  
ACCESSION BX112453  
VERSION BX112453.1 GI:27837617  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 642)

AUTHORS Ebert, L., Hell, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.  
TITLE Human Unigeneset - RZPD3  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGE998231165.  
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No. 972)  
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de  
This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r, Primer sequence: TTTCACACAGAAACAGCTATGAC.  
Location/Qualifiers

FEATURES  
source

1..642  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE998231165; IMAGE:1257620"  
/tissue\_type="pooled germ cell tumors"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP GC4"  
/note="Vector: pTZ19-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from 3 pooled  
germ cell tumors, and was then primed with a Not I +  
Eco RI primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pTZ19  
vector. Library is normalized. Library was constructed by  
Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 22.3%; Score 454; DB 5; Length 642;  
Best Local Similarity 99.7%; Pred. No. 5.3e-217;  
Matches 624; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

1407 GTTACAGAACTGGCTCTGTGTCGCGGTGGGGGTGTGCGGAATGTGACGACGTGC 1466  
DB 1 GTTACAGAACTGGCTCTGTGTCGCGGTGGGGGTGTGCGGAATGTGACGACGTGC 60  
1467 TGGGTGATCACTGCGCGCGGTGCTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1526  
DB 61 TGGGTGATCACTGCGCGCGGTGCTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
1527 CCCGCGCGGAG 1586  
DB 121 CCCGCGCGGAG 180  
1587 TGGG-GGGGGTGTGGCCCCCAGGCCCGCGCTGCGCCCTTGGGCTTGGCAAGATGAC 1645  
DB 181 TGGGAGGGGGTGTGGCCCCCAGGCCCGCGCTGCGCCCTTGGGCTTGGCAAGATGAC 240  
1646 ACTGCAGTCAAG 1705  
DB 241 ACTGCAGTCAAG 300  
1706 ACCTTGATGAGATCTGAG 1765  
DB 301 ACCTTGATGAGATCTGAG 360  
1766 CCTCTTGAACCCCAATATGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1825  
DB 361 CCTCTTGAACCCCAATATGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
1826 CCTCTTCTCAGTCTTGAAG 1885

Db 421 CCTCTCTCTCACTGCTGGAGCCGGCTGGATCAAGAGGGAGCCACCTTC 480  
Qy 1886 TTGTCACTGCTGGCTGTAAACAGCTCTGTCTTCTGGGAGACACGACCATCATGTGCTT 1945  
Db 481 TTGTCACTGCTGGCTGTAAACAGCTCTGTCTTCTGGGAGACACGACCATCATGTGCTT 540  
Qy 1946 GGAATTTAAACCTGCGCCCACTTCTACTCTGGAAGTCCCGGAGACCTCTCTGCTT 2005  
Db 541 GGAATTTAAACCTGCGCCCACTTCTACTCTGGAAGTCCCGGAGACCTCTCTGCTT 600  
Qy 2006 GGTGACCTACTATAAATATATAAATT 2031  
Db 601 GGTGACCTACTATAAATATATAAATT 626  
RESULT 3  
AY419551 1474 bp DNA linear GSS 17-DEC-2003  
LOCUS Pan troglodytes AIRE gene, VIRUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY419551  
VERSION AY419551.1 GI:39775508  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1474)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT These sequences were made by sequencing genomic exons and ordering  
them based on alignment  
FEATURES  
source location/Qualifiers  
1..1474  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>1474  
/gene="AIRE"  
/locus\_tag="HC6928"  
ORIGIN  
Query Match 11.2%; Score 229; DB 9; Length 1474;  
Best Local Similarity 99.4%; Pred. No. 1.5e-103;  
Matches 329; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 269 GAGAGCCTTATCTGAAGAAAAGAGGCTGCCCCAGGCTTCCAGCCCTCTGTCC 328  
Db 1 GAGAGCCTTATCTGAAGAAAAGAGGCTGCCCCAGGCTTCCAGCCCTCTGTCC 60  
Qy 329 TGGCTGCTGACCCAGAGCTTCCAGGCTTCTGAGGCTGCTTCAAGAGC 388  
Db 61 TGGCTGCTGACCCAGAGGCTTCCAGGCTTCTGAGGCTGCTTCAAGAGC 120  
Qy 389 TACAACTGAGAGGCTATGAGCGGCTGAGCCCATCTGAGAGAGCTTCCCAAAGTGG 448  
Db 121 TACAACTGAGAGGCTATGAGCGGCTGAGCCCATCTGAGAGAGCTTCCCAAAGTGG 180  
Qy 449 GACCTCAGCCAGCCCGGAGAGGGAGAGCCCGCGCGCTCCCAAGGCTTGTACCG 508

Db 181 GACCTCAGCCAGCCCGGAGAGGAGAGAGCCCGCGCGCTCCCAAGGCTTGTACCG 240  
Qy 509 CCACCCAGACTCCCAACCAAGAGAGGCTTCAAGAGGCTGAGCTGCCGCGCAGCA 568  
Db 241 CCACCCAGACTCCCAACCAAGAGAGGCTTCAAGAGGCTGAGCTGCCGCGCAGCA 300  
Qy 569 GCCCTGACTCCCAAGAGGAGCGCCGAGCCGAG 599  
Db 301 GCCCTGACTCCCAAGAGGAGCGCCGAGCCGAG 331  
RESULT 4  
AA742555/c 433 bp mRNA linear EST 22-JAN-1998  
LOCUS nx30c11.81 NCI\_CGAP\_GC4 Homo sapiens cDNA clone IMAGE:1257620 3',  
DEFINITION mRNA sequence.  
ACCESSION AA742555  
VERSION AA742555.1 GI:2782137  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-rc@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILMIL at:  
[www.bio.lnlnl.gov/bbrp/image/image.html](http://www.bio.lnlnl.gov/bbrp/image/image.html)  
Insert Length: 701 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence etop: 432.  
FEATURES  
source location/Qualifiers  
1..433  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1257620"  
/tissue="type="pooled germ cell tumors"  
/lab\_host="DH10B"  
/clone\_id="NCI CGAP GC4"  
/note="Vector: pRTT3-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from 3 pooled  
germ cell tumors, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pRTT3  
vector. Library is normalized. Library was constructed by  
Bento Soares and M. Fatima Bonaldo."  
ORIGIN  
Query Match 7.7%; Score 157; DB 1; Length 433;  
Best Local Similarity 99.6%; Pred. No. 3.1e-67;  
Matches 277; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1754 GCGGCCCTCTCTCTCTGACCCCAAGATGCGCGGAGCATGCACTGTAGAGAGTGG 1813  
Db 279 GCGGCCCTCTCTCTCTGACCCCAAGATGCGCGGAGCATGCACTGTAGAGAGTGG 220  
Qy 1814 CTGAGAAAGACACTCTTCTCTCACTCTGGAAGCGCGCGGCTGGAGTCAAGAGGGGA 1873  
Db 219 CTGAGAAAGACACTCTTCTCTCACTCTGGAAGCGCGCGGCTGGAGTCAAGAGGGGA 160



QY 1874 CAGGCGACCTCTTGTGAGTGGCTGTGAACAGCTGTGTTCTGGGAGACACGAGC 1933  
|||||  
Db 159 CAGGCGACCTCTTGTGAGTGGCTGTGAACAGCTGTGTTCTGGGAGACACGAGC 101  
|||||  
QY 1934 CATCATGTGCTGGAATTAACCCCTGCCCCACTTCTACTGTGAAAGTCCCCGGAGC 1993  
|||||  
Db 100 CATCATGTGCTGGAATTAACCCCTGCCCCACTTCTACTGTGAAAGTCCCCGGAGC 41  
|||||  
QY 1994 CTCTCTTGGCTGTGAGCTACTTAAATTAATAAATT 2031  
|||||  
Db 40 CTCTCTTGGCTGTGAGCTACTTAAATTAATAAATT 3  
|||||  
RESULT 5  
HSDML04M6 200 bp DNA linear GSS 17-AUG-1998  
LOCUS Homo sapiens DNA for trapped exon (ID MDL04M06), genomic survey  
DEFINITION  
ACCESSION AJ001114  
VERSION AJ001114.1 GI:3426119  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Nagamine, K., Peterson, P., Scott, H.S., Kudoh, J., Minoshima, S.,  
Heino, M., Krohn, K.J., Lalitoti, M.D., Myllis, P.E., Antonarakis, S.E.,  
Kawasaki, K., Asakawa, S., Ito, F. and Shimizu, N.  
TITLE Mutations in a novel zinc finger protein AIR are responsible for  
autoimmune polyglandular disease type I (APCED)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 200)  
AUTHORS Scott, H.S.  
TITLE Direct Submission  
JOURNAL Submitted (14-AUG-1997) Scott H.S., Department of Genetics and  
Microbiology, University of Geneva Medical School, 1 rue Michel  
Servet, Geneva, 1211, SWITZERLAND  
FEATURES  
source  
1..200  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="q22.3"  
1..200  
/note="trapped exon"  
ORIGIN  
exon  
Query Match 3.4%; Score 69; DB 9; Length 200;  
Best Local Similarity 100.0%; Pred. No. 7.6e-23;  
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 884 AGTGGCCGAGAGCTGTGGTTCAGAGGAGGAGCCAGGAGGCTGCCCGGTGAGGT 943  
|||||  
Db 96 AGTGGCCGAGAGCTGTGGTTCAGAGGAGGAGCCAGGAGGCTGCCCGGTGAGGT 155  
|||||  
QY 944 GAGGCTAGG 952  
|||||  
Db 156 GAGGCTAGG 164  
|||||  
RESULT 6  
AQ474914 350 bp DNA linear GSS 23-APR-1999  
LOCUS CITR1-B1-2591L22.TF CITR1-B1 Homo sapiens genomic clone 2591L22,  
DEFINITION genomic survey sequence.  
ACCESSION AQ474914  
VERSION AQ474914.1 GI:4657033  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 350)  
Zhao, S., Adams, M.D., Niernan, W., Malek, J., Shinya, H., Simon, M. and  
Venter, J.C.  
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready  
Map Building  
JOURNAL Unpublished (1997)  
COMMENT Other\_GSSs: CITR1-B1-2591L22.TF  
Contact: Shaying Zhao, William Niernan, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13-21  
Classes: BAC ends.  
FEATURES  
source  
1..350  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2591L22"  
/sex="male"  
/cell\_type="sperm"  
/clone\_lib="CITR1-B1"  
/note="Vector: pBeloBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
Caltech Human BAC Library p"  
ORIGIN  
Query Match 3.2%; Score 65; DB 8; Length 350;  
Best Local Similarity 100.0%; Pred. No. 7.8e-21;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1638 AGGATGACACTGCTCCAGTCAAGGAGGCTGTGCAAGAGGATGACTGAGTCCCTTGA 1697  
|||||  
Db 111 AGGATGACACTGCTCCAGTCAAGGAGGCTGTGCAAGAGGATGACTGAGTCCCTTGA 170  
|||||  
QY 1698 GCGAG 1702  
|||||  
Db 171 GCGAG 175  
|||||  
RESULT 7  
BE627856 451 bp mRNA linear EST 24-AUG-2000  
LOCUS BE627856  
DEFINITION u49c07.y1 Soares\_thymus\_2kbMT Mus musculus cDNA clone  
IMAGE:3375276 5' similar to TR:Q920E3 Q920E3 AIRE PROTEIN. ;, mRNA  
sequence.  
ACCESSION BE627856  
VERSION BE627856.1 GI:9908250  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 451)  
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c99b@rs-tnl.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:1084880  
Seq primer: -40RP from Gibco  
High quality sequence stop: 356.  
FEATURES  
source  
1..451  
Location/Qualifiers

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3375276"  
/sex="male"  
/tissue\_type="Thymus"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_thymus\_2NbMT"  
/note="vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTTCACATCTGAGTGGAGCGCGCTGCTTTTGTGTGT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through two  
rounds of normalization, and was constructed by Bento  
Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 1.6%; Score 32; DB 2; Length 451;  
Best Local Similarity 100.0%; Pred. No. 0.00034;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACATCGCGCTGCCACTTCCC 1515

DB 41 GCCGCTGCTTCACATCGCGCTGCCACTTCCC 72

RESULT 8  
AY419552 1512 bp DNA linear GSS 17-DEC-2003

DEFINITION Mus musculus AIRE gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

ACCESSION AY419552  
VERSION AY419552.1 GI:39775509

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1512)  
Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios

TITLE  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 1512)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE  
JOURNAL Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA

COMMENT  
These sequences were made by sequencing genomic exons and ordering  
them based on alignment.

FEATURES  
source location/Qualifiers

1..1512  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>1512  
/gene="AIRE"  
/locus\_tag="HGM6928"

## ORIGIN

Query Match 1.6%; Score 32; DB 9; Length 1512;

Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 GCCGCTGCTTCACATCGCGCTGCCACTTCCC 1515

DB 1213 GCCGCTGCTTCACATCGCGCTGCCACTTCCC 1244

## RESULT 9

AI552580 384 bp mRNA linear EST 23-MAR-1999  
LOCUS vx25b07.x1 Soares mammary\_gland NBMWG Mus musculus cDNA clone  
DEFINITION IMAGE:1265557 3', mRNA sequence.

ACCESSION AI552580  
VERSION AI552580.1 GI:4484943

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE  
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP);  
Tumor Gene Index  
JOURNAL Unpublished (1997)

COMMENT  
Email: cgapb-remail.nih.gov  
Contact: Robert Strausberg, Ph.D.  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:668109

This clone was previously sequenced on the 5' end only, this new  
data is from the 3' end  
High quality sequence stop: 381.

## FEATURES

source location/Qualifiers

1..384  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1265557"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares mammary gland NBMWG"  
/note="Organ: mammary gland; Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5',  
TGTTCACATCTGAGTGGAGCGCGCTGCTTTTGTGTGT  
3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

## ORIGIN

Query Match 1.5%; Score 30; DB 1; Length 384;  
Best Local Similarity 100.0%; Pred. No. 0.0035;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 GGCACTCTGAGTGGCCATCGAGCAGT 1744

DB 296 GGCACTCTGAGTGGCCATCGAGCAGT 267

## RESULT 10

CG513175 420 bp DNA linear GSS 01-OCT-2003  
LOCUS OST66421 Mus musculus 129SV/Ev Mus musculus genomic clone OST66421,  
DEFINITION genomic survey sequence.  
ACCESSION CG513175

VERSION CG513175.1 GI:37299748  
KEYWORDS CSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 420)  
AUTHORS Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Figgott, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A., Fridele, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C., Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Slichtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.  
TITLE Mx1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)  
COMMENT Contact: Zambrowicz BP  
Omnibank  
Lexicon Genetics Incorporated  
4000 Research Forest Drive, The Woodlands, TX 77381, USA  
Email: materials@lexgen.com  
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)  
Class: Gene Trap.  
ORIGIN  
Location/Qualifiers  
1..420  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129Sv/Ev"  
/db\_xref="taxon:10090"  
/clone="OST66421"  
/cell\_type="embryonic stem cell"  
/clone\_1ib="Mus musculus 129Sv/Ev"  
Query Match 1.5%; Score 30; DB 9; Length 420;  
Best Local Similarity 100.0%; Pred.No. 0.0034;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1715 GGCATCTCGACGTGGCCATCCAGACGATG 1744  
|||||  
170 GGCATCTCGACGTGGCCATCCAGACGATG 199  
RESULT 11  
BB630816/c 511 bp mRNA linear EST 25-AUG-2000  
LOCUS uu49c07.x1 Soares thymus 2NBMT Mus musculus cDNA clone  
DEFINITION IMAGE:3375276 3' similar to TR:Q920E3 Q920E3 AIRE PROTEIN.; mRNA  
sequence.  
ACCESSION BE630816  
VERSION BE630816  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 511)  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
COMMENT Tumor Gene Index  
Unpublished (1997)  
Other ESTs: uu49c07.y1  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:1084880  
High quality sequence stop: 391.  
Location/Qualifiers  
1..511  
/organism="Mus musculus"

ORIGIN  
Query Match 1.5%; Score 30; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred.No. 0.0034;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1715 GGCATCTCGACGTGGCCATCCAGACGATG 1744  
|||||  
282 GGCATCTCGACGTGGCCATCCAGACGATG 253  
RESULT 12  
AA703834 459 bp mRNA linear EST 24-DEC-1997  
LOCUS ag78g12.x1 Stratagene hnt neuron (#937233) Homo sapiens cDNA clone  
DEFINITION IMAGE:1140646 5' similar to contains Alu repetitive  
element; contains element OPR repetitive element.; mRNA sequence.  
ACCESSION AA703834  
VERSION AA703834.1 GI:2713752  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 459)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepcie, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28m13 rev1 ET from Amersham.  
Location/Qualifiers  
1..459  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1140646"  
/dev\_stage="hnt neurons"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_1ib="Stratagene hnt neuron (#937233)"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI; Cloned unidirectionally. Primer: Oligo dT.  
Differentiated, post mitotic hnt neurons. Average insert  
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'  
GAATTGGGACGACG 3' -3' adaptor sequence: 5'  
GAAATTCGACGACG 3' -3' adaptor sequence: 5'

```

ORIGIN      CTCGAGTGTGTTTTTTTTTTTTTTT 3''

Query Match      1.3%; Score 26; DB 1; Length 459;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2011 CCTACTAAAAATATATAAATTAGCTG 2036
      |||
      384 CCTACTAAAAATATATAAATTAGCTG 409

RESULT 13
BZ277407      478 bp      DNA      linear      GSS 15-OCT-2002
LOCUS      CH230-392014.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION      CH230-392014, genomic survey sequence.
ACCESSION      BZ277407
VERSION      BZ277407
KEYWORDS      GSS.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 478)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-392014.TV
COMMENT      Contact: Shaying Zhao
      Department of Eukaryotic Genomics
      The Institute for Genomic Research
      9712 Medical Center Dr., Rockville, MD 20850, USA
      Tel: 301 838 0200
      Fax: 301 838 0208
      Email: szhao@tigr.org
      Clones are derived from the rat BAC library CHORI-230
      (http://www.chori.org/bacpac/rat230.htm). For BAC library
      availability, please contact Pieter de Jong (pjejong@mail.cho.org).
      Clones may be purchased from BACPAC Resources.
      (http://www.chori.org/bacpac/orering/information.htm). BAC end
      page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
      Place: 392 row: 0 column: 14
      Seq primer: T7
      Class: BAC ends.
      Location/Qualifiers
        1..478
          /organism="Rattus norvegicus"
          /mol_type="genomic DNA"
          /strain="BN/SsNHsd/MCW"
          /db_xref="taxon:10116"
          /clone="CH230-392014"
          /sex="Female"
          /cell_type="Brain"
          /clone_11b="CHORI-230 Segment 2"
          /note="Vector: PTARBAC1.3; Site_1: MboI; Site_2: MboI;
          CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
          Pieter de Jong"

ORIGIN

Query Match      1.3%; Score 26; DB 8; Length 478;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      296 GGCTGCCCGAGGCTTCCACGCCCT 321
      |||
      415 GGCTGCCCGAGGCTTCCACGCCCT 440

Db

RESULT 14

```

```

AQ239234
LOCUS      AQ239234      479 bp      DNA      linear      GSS 21-APR-1999
DEFINITION      RPC111-67M12.TK RPC1-11 Homo sapiens genomic clone RPC1-11-67M12,
      genomic survey sequence.
ACCESSION      AQ239234
VERSION      AQ239234.1 GI:3671525
KEYWORDS      GSS.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 479)
Adams,M.D., Rounley,S.D., Zhao,S., Base,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1999)
Other GSSs: RPC111-67M12.TV
COMMENT      Contact: Mark Adams
      Department of Eukaryotic Genomics
      The Institute for Genomic Research
      9712 Medical Center Dr., Rockville, MD 20850, USA
      Tel: 301 838 0200
      Fax: 301 838 0208
      Email: mdamas@tigr.org
      Clones are derived from the human BAC library RPC1-11. For BAC
      library availability, please contact Pieter de Jong
      (pjejong@tigr.med.buffalo.edu). Clones may be purchased from
      BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/) or from
      Research Genetics (info@resgen.com). BAC end search page:
      http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
      Seq primer: T7
      Class: BAC ends.
      Location/Qualifiers
        1..479
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="GDB:7525643"
          /db_xref="taxon:9606"
          /clone="RPC1-11-67M12"
          /sex="Male"
          /cell_type="Lymphocytes"
          /clone_11b="RPC1-11"
          /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
          RPC11 Human Male BAC Library"

ORIGIN

Query Match      1.3%; Score 26; DB 8; Length 479;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2011 CCTACTAAAAATATATAAATTAGCTG 2036
      |||
      223 CCTACTAAAAATATATAAATTAGCTG 248

RESULT 15
CG585187      512 bp      DNA      linear      GSS 02-OCT-2003
LOCUS      OGT323246 Mus musculus 129Sv/Ev Mus musculus genomic clone
DEFINITION      OGT323246, genomic survey sequence.
ACCESSION      CG585187
VERSION      CG585187
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 512)
Zambrowicz,B.P., Abuln,A., Ramirez-Solis,R., Richter,L.J.,
Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jiang,C.,
Key,B.W., Jr., Klipp,P., Kohhauf,B., Ma,Z.-Q., Matkeisch,D.,
Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,

```

**TITLE**  
Zhu O., Person C. and Sands A.T.  
Mnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

**COMMENT**  
Contact: Zambrowicz BP  
OmniBank  
Lexicon Genetics Incorporated  
4000 Research Forest Drive, The Woodlands, TX 77381, USA  
Email: materials@lexgen.com  
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6676):608-11)  
Class: Gene Trap.

**FEATURES**  
source  
Location/Qualifiers  
1..512  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129Sv/Ev"  
/db\_xref="taxon:10090"  
/clone="OST323246"  
/cell\_type="embryonic stem cell"  
/clone\_lib="Mus musculus 129Sv/Ev"

**ORIGIN**

Query Match  
Best Local Similarity 1.3%; Score 26; DB 9; Length 512;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy**  
296 GGCTGCCCCAGGCTTCACAGCCCT 321  
|||||  
268 GGCTGCCCCAGGCTTCACAGCCCT 293

**Db**  
296 GGCTGCCCCAGGCTTCACAGCCCT 321  
|||||  
268 GGCTGCCCCAGGCTTCACAGCCCT 293

**RESULT 16**  
BY745644 632 bp mRNA linear EST 17-DEC-2002  
LOCUS BY745644 RIKEN full-length enriched, 3 days neonate thymus Mus  
DEFINITION musculus cDNA clone A630047M03 5', mRNA sequence.  
ACCESSION BY745644  
VERSION BY745644.1 GI:27172953  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

**REFERENCE**  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 632)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bull, C., Hume, D.A., Quackenbush, J., Schiml, L.M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T., Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S., Guerinot, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurahachi, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglocz, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shinada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Welle, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kobayashi, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

**TITLE**  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanishi, A., Yamamatsu, M. and Hayashizaki, Y.  
Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichipillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

**FEATURES**  
source

Location/Qualifiers  
1..632  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="A630047M03"  
/tissue\_type="thymus"  
/dev\_stage="3 days neonate"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 3 days neonate thymus"  
/note="Site 1: SalI, Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3']. cDNA was prepared by using trihaloase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTTCGATTAATTAAATATCCCTCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified plasmid pUC19 (+) after bulk excision from Lambda

**ORIGIN**



(E-mail:chimbpe@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
PRIMERS

## Sequencing: T7

## LIBRARY

Vector : pBACE3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

Location/Qualifiers

1..715

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

/clone="RP43-006P12.T7"

/sex="male"

/cell\_type="lymphocytes"

/clone\_11b="RPCI-43 Chimpanzee Male BAC Library"

1.3%; Score 26; DB 9; Length 715;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2011 CCTACTAAATATATAAATTAGCTG 2036

333 CCTACTAAATATATAAATTAGCTG 308

## RESULT 20

## LOCUS

CN315153 742 bp mRNA linear EST 16-MAY-2004

DEFINITION 17000418008497 GRN\_ES Homo sapiens CDNA 5', mRNA sequence.

ACCESSION CN315153

VERSION CN315153.1 GI:47331567

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 742)

Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,

Li,Y., Xu,C., Fang,R., Guegler,K., Rao,W.S., Mandalam,R.,

Lebowksi,J. and Stanton,L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6): 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@genon.com

Insert Length: 742 Std Error: 0.00.

Location/Qualifiers

1..742

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="embryonic stem cells, cell lines H1, H7, and

H9"

/clone\_11b="GRN ES"

/note="oligo dt primed, full-length enriched cDNA library

from undifferentiated hES cell lines H1 (p32), H7 (p29),

and H9 (p26) maintained in feeder-free conditions"

QY 2011 CCTACTAAATATATAAATTAGCTG 2036  
 DB 234 CCTACTAAATATATAAATTAGCTG 259

## RESULT 21

## LOCUS

BQ431008 923 bp mRNA linear EST 24-MAY-2002

DEFINITION AGENCOURT\_7802230 NIH\_MGC\_92 Homo sapiens CDNA clone IMAGE:6042109

5', mRNA sequence.

ACCESSION BQ431008

VERSION BQ431008.1 GI:21170084

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 923)

NIH-MGC http://mhc.nci.nih.gov/.

Unpublished (1999)

National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.W.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be

found through the I.W.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LLM13280 row: p column: 14

High quality sequence stop: 584.

Location/Qualifiers

1..923

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_11b="NIH MGC 92"

/tissue\_type="embryonal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pCMV-Sport6; Site 1: NotI;

Site 2: SalI; Cloned, unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_MGC Library."

QY 2011 CCTACTAAATATATAAATTAGCTG 2036

DB 73 CCTACTAAATATATAAATTAGCTG 98

Query Match 1.3%; Score 26; DB 5; Length 923;

Best Local Similarity 100.0%; Pred. No. 0.35; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2011 CCTACTAAATATATAAATTAGCTG 2036

73 CCTACTAAATATATAAATTAGCTG 98

RESULT 22

AOS30389 119 bp DNA linear GSS 18-MAY-1999

DEFINITION RPCI-11-353L4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-353L4,

genomic survey sequence.

ACCESSION AOS30389

VERSION AOS30389.1 GI:4842432

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 119)

Zhao,S., Adams,W.D., Nieman,W., Malek,J., de Jong,P. and

Venter,J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
JOURNAL Map Building  
COMMENT Unpublished (1997)  
Contact: Shaying Zhao, William Niemman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieder@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genet cs (<http://inforesgen.com>). BAC end search page:  
[http://www.tigr.org/cdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html).  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..119  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7635435"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-35314"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_11b="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

ORIGIN  
Query Match 1.2%; Score 25; DB 8; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
|||||  
76 CTACTAAAAATATATAAATTAGCTG 100

RESULT 23 138 bp DNA linear GSS 19-MAY-1999  
AQ541617  
LOCUS RPCI-11-358M18.TV RPCI-11 Homo sapiens genomic clone  
DEFINITION RPCI-11-358M18, genomic survey sequence.  
ACCESSION AQ541617  
VERSION AQ541617.1 GI:4872073  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 138)  
Zhao, S., Adams, M.D., Niemman, W., Malek, J., de Jong, P. and  
Venter, J.C.  
Title Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building (1997)  
Unpublished (1997)  
Contact: Shaying Zhao, William Niemman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieder@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genet cs (<http://inforesgen.com>). BAC end search page:  
[http://www.tigr.org/cdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html).  
Seq primer: T7

JOURNAL  
COMMENT

Class: BAC ends.  
FEATURES  
source Location/Qualifiers  
1..138  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7637393"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-358M18"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_11b="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

ORIGIN  
Query Match 1.2%; Score 25; DB 8; Length 138;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
|||||  
110 CTACTAAAAATATATAAATTAGCTG 134

RESULT 24 149 bp mRNA linear EST 13-JAN-2001  
BF843049/c  
LOCUS R5-HT1036-141200-011-B09 HT1036 Homo sapiens CDNA, mRNA sequence.  
DEFINITION BF843049  
ACCESSION BF843049  
VERSION BF843049.1 GI:12197400  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 149)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.P., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/ICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?1-R5c5&2-R5-HT1036-141200-011-B09&3=2000-12-14&4=1>)  
Seq primer: puc 18 forward  
High quality sequence stop: 149.  
Location/Qualifiers  
1..149  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_11b="HT1036"  
/note="Organ: head neck; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORFESTS PCR (U.S. Letters patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of



ORIGIN tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 1.2%; Score 25; DB 2; Length 149;  
Beet Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
|||||  
Db 86 CTACTAAAAATATATAAATTAGCTG 62

## RESULT 25

LOCUS T70055 150 bp mRNA linear EST 23-FEB-1995  
DEFINITION ycl17c08.g1 Stragene lung (#937210) Homo sapiens cDNA clone  
IMAGE:80942 3' similar to contains Alu repetitive element; mRNA  
sequence.

ACCESSION T70055 GI:681203  
VERSION T70055.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 150)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,  
Chasse, S., Dietrich, N., Dubuque, T., Favello, A., Gish, M.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
MEDLINE 97044478  
PUBMED 8889549

## COMMENT

Contact: Wilson R.  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 222

Source: IMAGE Consortium, LNL This clone is available royalty-free  
through LNL; contact the IMAGE Consortium (info@image.llnl.gov)  
for further information.  
Insert Length: 222 Std Error: 0.00  
Seq primer: -21m13

High quality sequence stop: 476.

## FEATURES

source  
1..150  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="CDB:484559"  
/db\_xref="taxon:9606"  
/clone="IMAGE:80942"  
/sex="male"  
/dev\_stage="72 years"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/clone\_lib="Stragene lung (#937210)"  
/note="Organ: lung; Vector: pBluescript SK-; Site\_1:  
Scor1; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. normal lung. Average insert size: 1.0 kb;  
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGACGAG  
3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

## ORIGIN

Query Match 1.2%; Score 25; DB 7; Length 150;  
Beet Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
|||||  
Db 16 CTACTAAAAATATATAAATTAGCTG 40

## RESULT 26

LOCUS AW997297 152 bp mRNA linear EST 05-JUN-2000  
DEFINITION RC2-BN0048-150400-017-c10 BN0048 Homo sapiens cDNA, mRNA sequence.  
AM997297  
ACCESSION AW997297  
VERSION AW997297.1 GI:8257531  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 152)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Britones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the RAPSP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=RC2-BN0048-150  
400-017-c10&c3=2000-04-15&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 7  
High quality sequence stop: 152.

Location/Qualifiers

source  
1..152  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BN0048"  
/note="Organ: breast, normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 152;  
Beet Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
|||||  
Db 103 CTACTAAAAATATATAAATTAGCTG 127

## RESULT 27

LOCUS T70121 162 bp mRNA linear EST 23-FEB-1995  
DEFINITION ycl17c08.r1 Stragene lung (#937210) Homo sapiens cDNA clone  
IMAGE:80942 5' similar to contains Alu repetitive element; contains

MER22 repetitive element ;, mRNA sequence.

ACCESSION T70121  
 VERSION T70121.1 GI:681269  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 AUTHORS 1 (bases 1 to 162)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Maritz, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.,  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT 8889549  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 222  
 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
 Insert Length: 222 Std Error: 0.00  
 Seq primer: M13RP1  
 High quality sequence stop: 327.

FEATURES  
 source  
 Location/Qualifiers  
 1..162  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:484559"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:80942"  
 /sex="male"  
 /dev\_stage="72 years"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Stratagene lung (#937210)"  
 /note="Organ: lung; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT, normal lung. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."

ORIGIN

Query Match 1.2%; Score 25; DB 7; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 |||||  
 158 CTACTAAATATATAAATTAGCTG 134

RESULT 28  
 BF915807/c 179 bp mRNA linear EST 18-JUN-2001  
 LOCUS IL3-UT0114-081200-366-E12 UT0114 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION  
 ACCESSION BF915807  
 VERSION BF915807.1 GI:12307265  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 179)  
 Dias Neto, F., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsumura, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT 20202663  
 10737880  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=IL3&l2=IL3-UT0114-081200-366-E12&l3=2000-12-08&l4=1)  
 Seq primer: pUC 18 forward  
 High quality sequence start: 9  
 High quality sequence stop: 179.

FEATURES  
 source  
 Location/Qualifiers  
 1..179  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="UT0114"  
 /note="Organ: uterus; tumor; Vector: pUC18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 |||||  
 134 CTACTAAATATATAAATTAGCTG 110

RESULT 29  
 AA664564/c 184 bp mRNA linear EST 13-FEB-1998  
 LOCUS nu71d07.s1 NCI\_CGAP Alvi Homo sapiens cDNA clone IMAGE:1216141  
 DEFINITION similar to contains Alu repetitive element; contains element MSRI  
 MSRI repetitive element ;, mRNA sequence.  
 ACCESSION AA664564  
 VERSION AA664564.1 GI:2619177  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 184)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgs@bbs-remail.nih.gov  
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.

Query Match	1.2%	Score 25;	DB 2;	Length 190;
Best Local Similarity	100.0%;	Pred. No. 1.2;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	2012	CTACTAAATAATATTTAAATTGCTG	2036	

Db 62 CTAATAATAATAATAATTAGCTG 86

RESULT 32  
LOCUS BF883245/c 194 bp mRNA linear EST 17-JAN-2001  
DEFINITION QV3-ET0174-081200-508-b04 ET0174 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF883245  
VERSION BF883245.1 GI:12273371  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV3&t=QV3-ET0174-081200-508-b04&t3=2000-12-08&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 194.  
Location/Qualifiers  
1..194  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="ET0174"  
/note="Organ: lung; tumor; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 194;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAATAATAATAATTAGCTG 2036  
|||||  
102 CTAATAATAATAATAATTAGCTG 78

RESULT 33  
LOCUS A1193409/c 200 bp mRNA linear EST 29-OCT-1998  
DEFINITION G564G03.x1 Soares fetal\_lung\_NbH19W Homo sapiens cDNA clone  
IMAGE:1743172 3' similar to contains Alu repetitive element;; mRNA  
sequence.  
ACCESSION A1193409

VERSION A1193409.1 GI:3744618  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-r@mail.nih.gov  
This clone is available royally-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 605 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 189.  
Location/Qualifiers  
1..200  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:1743172"  
/dev\_stage="19 weeks"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal lung NbH19W"  
/note="Organ: lung; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCCGCAATTTTCTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbH19W."

ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 200;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAATAATAATAATTAGCTG 2036  
|||||  
189 CTAATAATAATAATAATTAGCTG 165

RESULT 34  
LOCUS AA626595  
DEFINITION AA626595 211 bp mRNA linear EST 06-MAR-1998  
IMAGE:844660 5' similar to contains Alu repetitive element; contains TAR1.c3 TAR1 repetitive element;; mRNA sequence.  
ACCESSION AA626595  
VERSION AA626595.1 GI:2538982  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Matra,M., Martin,J., Moore,B., Scheilenberg,K., Stepcoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
TITLE Wabhu-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estwason.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 375 Std Error: 0.00  
 Seq primer: -28nt3 rev1 ET from Amerham  
 High quality sequence stop: 152.

## FEATURES

source

Location/Qualifiers  
 1. 211  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:944660"  
 /issue\_type="lung carcinoma"  
 /cell\_line="NCI-H69"  
 /dev\_stage="cell line NCI-H69"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone\_lib="Stratagene lung carcinoma 937218"  
 /note="Organ: lung; Vector: pBluescript SK-; Site 1:  
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
 Oligo dt. Small cell carcinoma cell line NCI-H69. Average  
 insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
 sequence: 5' GAATTCGGCAGCG 3' -3' adaptor sequence: 5'  
 CTCGAGTTT TTT TTT TTT TTT TTT TTT 3'"

## ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 |||  
 118 CTACTAAATATATAAATTAGCTG 142

RESULT 35  
 BI007053 232 bp mRNA linear EST 13-JUN-2001  
 LOCUS RCS-RT0054-170101-021-H09 RT0054 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BI007053  
 ACCESSION BI007053.1 GI:14411124  
 VERSION EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 232)  
 Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PABESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RCS&t2=RCS-RT0054-  
 170101-021-H09&t3=2001-01-17&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 51

High quality sequence stop: 232.  
 Location/Qualifiers  
 1. 232  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"

## FEATURES

source

/clone\_lib="R10054"  
 /note="Organ: kidney tumor; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Query Match 1.2%; Score 25; DB 4; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 |||  
 133 CTACTAAATATATAAATTAGCTG 157

RESULT 36  
 BG615666 235 bp mRNA linear EST 18-APR-2001  
 LOCUS 602643025P1 NIH\_MGC\_61 Homo sapiens cDNA clone IMAGE:4773799 5',  
 DEFINITION BG615666  
 mRNA sequence.  
 ACCESSION BG615666.1 GI:13667037  
 VERSION EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 235)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: CLONETECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L16M1646 row: b column: 08  
 High quality sequence stop: 234.

## FEATURES

source

Location/Qualifiers  
 1. 235  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4773799"  
 /tissue\_type="embryonal carcinoma"  
 /lab\_host="RDH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_61"  
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgccgctggc); Site 2: SfiI (ggccatagggc);  
 Double-stranded cDNA was prepared from cell line RNA. 5'  
 and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CACGGCATTATGCG-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCCGACGCGCGGCGATG-dT(30)BN-3'  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC



REFERENCE 1 (bases 1 to 248)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncigap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLNT at: [www.bio.linn.gov/bbrp/image/image.html](http://www.bio.linn.gov/bbrp/image/image.html)  
 Seq primer: -40m3 fwd, BT from Amerham.  
 Location/Qualifiers  
 1..248  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1386281"  
 /sex="female"  
 /tissue\_type="ovary"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI-CGAP OV2"  
 /note="Vector: PAMP10; mRNA made from invasive ovarian tumor, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

ORIGIN  
 Query Match 1.2%; Score 25; DB 1; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATTAATAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 Db 79 CTACTAAATATTAATAATTAGCTG 103

RESULT 40  
 BE2975 252 bp DNA linear GSS 20-JUN-1998  
 LOCUS CIT-HSP-2007P8.TF CIT-HSP Homo sapiens genomic clone 2007P8,  
 genomic survey sequence.  
 ACCESSION B52975  
 VERSION B52975.1 GI:2607309  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 252)  
 Adams,M.D., Rounsley,S.D., Field,C.E., Baas,S., Linher,K.,  
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
 Simon,M. and Venter,J.C.  
 Use of a random BAC End Sequence Database for Sequence-Ready Map  
 Building  
 Unpublished (1997)  
 Other\_GSSs: CIT-HSP-2007P8.TRC  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: maddams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
[http://www.tigr.org/cdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html)

Seq primer: M13-21  
 Class: BAC ends.  
 Location/Qualifiers  
 1..252  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7040242"  
 /db\_xref="taxon:9606"  
 /clone="2007P8"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /clone\_lib="CIT-HSP"  
 /note="Vector: pBSL0BAC11; Site\_1: HindIII; Site\_2: HindIII"

ORIGIN  
 Query Match 1.2%; Score 25; DB 8; Length 252;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATTAATAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 Db 210 CTACTAAATATTAATAATTAGCTG 234

RESULT 41  
 BF851782 255 bp mRNA linear EST 16-JUN-2001  
 LOCUS BF851782  
 DEFINITION CM3-EN0079-051200-521-c04 EN0079 Homo sapiens CDNA, mRNA sequence.  
 ACCESSION BF851782  
 VERSION BF851782.1 GI:12238944  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 255)  
 Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM3&e2=CM3-EN0079-051200-521-c04&t3=2000-12-05&t4=1>  
 Seq primer: puc 18 forward  
 High quality sequence start: 17  
 High quality sequence stop: 255.  
 Location/Qualifiers  
 1..255  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="EN0079"  
 /note="Organ: Lung normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORBSTES PCR (U.S. Letters Patent application

No. 196, 716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 255;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036

Db 121 CTACTAAAAATATATAAATTAGCTG 97

## RESULT 42

AA618344 256 bp mRNA linear EST 21-OCT-1997  
LOCUS nql8a03.s1 NCI CGAP Thy1 Homo sapiens cDNA clone IMAGE:1144204  
DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
ACCESSION AA618344  
VERSION AA618344.1 GI:2505549  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 256)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert length: 348 Std Error: 0.00  
Seq primer: -40m3 fwd. ET from Amersham.

## FEATURES

source

Location/Qualifiers  
1..256  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1144204"  
/tissue\_type="thyroid"  
/lab\_host="DH10B"  
/clone\_1ib="NCI CGAP Thy1"  
/note="Vector: PAMP10; mRNA made from invasive thyroid  
tumor, cDNA made by oligo-dT priming. Non-directionally  
cloned. Size-selected on agarose gel, average insert size  
600 bp. Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."

## ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036

Db 174 CTACTAAAAATATATAAATTAGCTG 150

## RESULT 43

F26713 256 bp mRNA linear EST 13-MAY-1999  
LOCUS F26713  
DEFINITION HSPD14290 HM3 Homo sapiens cDNA clone s4000059B10, mRNA sequence.

ACCESSION F26713  
VERSION F26713.1 GI:4812339  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 256)  
Lanfanchi, G., Moraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,  
Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.  
Identification of 4370 expressed sequence tags from a  
3'-end-specific cDNA library of human skeletal muscle by DNA  
sequencing and filter hybridization  
Genome Res. 6 (1), 35-42 (1996)

JOURNAL  
MEDLINE  
PUBMED  
96276048  
8681137

## COMMENT

Contact: Valle G.  
CIRI Biotechnology Centre  
University of Padua  
Via Trieste 75, 35121 Padua, Italy  
ABI Chromatograms and other information are available on WWW at  
<http://group.bio.unipd.it>.

## FEATURES

source

Location/Qualifiers  
1..256  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="s4000059B10"  
/sex="female"  
/tissue\_type="pectoral muscle (after mastectomy)"  
/clone\_1ib="HM3"  
/note="Vector: pCDNAII (Invitrogen); Site 1: BstXI;  
Site 2: NotI. The library is not subcloned nor normalized.  
Lanfanchi. This library is not subcloned nor normalized.  
The first strand cDNA was primed with a biotinylated  
oligo-dT-NotI primer  
(5'-biotin-AACCGGCTCGAGCGCGCTTTTCTTTTCTTTT-3'). The  
ds cDNA was sonicated and size-selected in the range  
350-550 bp. The 3' specific fragments were selected by  
streptavidin coated magnetic beads, ligated to  
non-palindromic BstXI adapters, NotI digested and  
directionally cloned into BstXI-NotI cut pCDNAII vector."

## ORIGIN

Query Match 1.2%; Score 25; DB 7; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036

Db 214 CTACTAAAAATATATAAATTAGCTG 238

## RESULT 44

N73518 256 bp mRNA linear EST 19-MAR-1996  
LOCUS 2449e03.s1 Soares fetal liver spleen INFIS Homo sapiens cDNA clone  
DEFINITION IMAGE:295900 3' similar to contains Alu repetitive element; mRNA  
sequence.

ACCESSION N73518  
VERSION N73518.1 GI:1230803  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 256)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Hillier, L., Clark, N., Dubuque, T., Ellison, K., Hawkins, M.,  
Holman, M., Hallman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevasakis, E., Waterston, R., Williamson, A., Woldmann, P. and  
Wilson, R.

TITLE The Washu-Merck EST Project



JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: m13 -40 forward  
High quality sequence stop: 220.

FEATURES  
source  
1. 256  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:1240821"  
/db\_xref="taxon:9606"  
/clone="IMAGE:295900"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_11b="Soares fetal liver spleen INFUS"  
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5': AACTGGAAGAATTATTAAGATCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN  
Query Match 1.2%; Score 25; DB 7; Length 256;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTTAAATTTAAATTTAGCTG 2036  
|||||  
165 CTACTTAAATTTAAATTTAGCTG 141

RESULT 45  
CD516253 260 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT.14372378 NIH\_MGC.181 Homo sapiens cDNA clone  
DEFINITION IMAGE:30394876 5', mRNA sequence.  
ACCESSION CD516253  
VERSION CD516253.1 GI:31447971  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 260)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDAM474 row: b column: 05  
High quality sequence start: 33  
High quality sequence stop: 260.

JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: m13 -40 forward  
High quality sequence stop: 220.

FEATURES  
source  
Location/Qualifiers  
1. 260  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30394876"  
/tissue\_type="White Matter"  
/dev\_stage="Unknown"  
/lab\_host="DH10B-Ton A (T1 and T5 phage resistances)"  
/clone\_11b="NIH\_MGC\_181"  
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV  
(destroyed); Library is oligo-dT primed and directionally  
cloned (EcoRV site is destroyed upon cloning). Average  
insert size 1.42 kb. Library was constructed by  
(Invitrogen). Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 1.2%; Score 25; DB 6; Length 260;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTTAAATTTAAATTTAGCTG 2036  
|||||  
77 CTACTTAAATTTAAATTTAGCTG 101

RESULT 46  
AA847504 266 bp mRNA linear EST 04-MAR-1998  
LOCUS oel9b03.s1 NCI CGAP OV2 Homo sapiens cDNA clone IMAGE:136317  
DEFINITION similar to contains Alu repetitive element; contains element MER22  
repetitive element; mRNA sequence.  
ACCESSION AA847504  
VERSION AA847504.1 GI:2934022  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 266)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bbbrp/image.html  
Seq primer: -40m13 fwd. ET from Amerham.

FEATURES  
source  
1. 266  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:136317"  
/sex="female"  
/tissue\_type="ovary"  
/lab\_host="DH10B"  
/clone\_11b="NCI CGAP OV2"  
/note="Vector: pAMP10; mRNA made from invasive ovarian  
tumor. cDNA made by oligo-dT priming. Non-directionally  
cloned. Site-selected on agarose gel, average insert size  
600 bp. Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."

ORIGIN  
Query Match 1.2%; Score 25; DB 1; Length 266;

Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036  
Db 76 CTACTAAAAATATATAAATTAGCTG 100

## RESULT 47

AA829490/c

LOCUS 270 bp mRNA linear EST 25-MAR-1998  
DEFINITION cd06904.s1 NCI CGAP GCBI Homo sapiens CDNA clone IMAGE:1358454 3'

ACCESSION AA829490  
VERSION AA829490.1 GI:2902589

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Louis M. Straud, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert length: 1200 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amerham

High quality sequence stop: 223.

## FEATURES

source

Location/Qualifiers

1..270

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1358454"

/tissue\_type="germinal center B cell"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP GCBI"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand CDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Straud (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was primed with a Not I - oligo (dT) primer

[5'-TGTTACCAATCTGAAGTGGAGCGCCGCTCATTTTCTTTTCTTTT-3'

] Double-stranded CDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 270;  
Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036  
Db 178 CTACTAAAAATATATAAATTAGCTG 154

RESULT 48

AA468966/c 272 bp mRNA linear EST 14-AUG-1997  
LOCUS nc79b09.r1 NCI CGAP Pr2 Homo sapiens CDNA clone IMAGE:783545  
DEFINITION similar to contains Alu repetitive element; mRNA sequence.

ACCESSION AA468966  
VERSION AA468966.1 GI:2195500

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquani, M.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Krizman, Ph.D.

CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert length: 534 Std Error: 0.00

Seq primer: -28ml3 rev1 ET from Amerham

High quality sequence stop: 255.

## FEATURES

source

Location/Qualifiers

1..272

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:783545"

/sex="Male"

/dev\_stage="45 years old"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Pr2"

/note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st strand CDNA was primed with oligo (dT) 17 on 50 ng of DNase-treated, total cellular RNA obtained from

5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial

neoplasia 2 (PIN2) cells. Double-stranded CDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the

CDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning

method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was

constructed by David Krizman."

## ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATATAAATTAGCTG 2036  
Db 167 CTACTAAAAATATATAAATTAGCTG 143

## RESULT 49

CD357167/c

LOCUS CD357167 279 bp mRNA linear EST 29-MAY-2003  
DEFINITION AGENCOURT\_14254892 NIH\_MGC\_187 Homo sapiens CDNA clone IMAGE:30404658 5', mRNA sequence.

ACCESSION CD357167  
VERSION CD357167.1 GI:31128578

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 279)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Place: NDCM190 row: 1 column: 19  
 High quality sequence stop: 257.  
 Location/Qualifiers  
 1..279  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30404658"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH-MGC-187"  
 /note="Organ: Blood vessels - aorta, basilar and artery; Vector: pDNR-LIB; Site 1: SfiI (ggcgatcgcc); Site 2: SfiI (ggcgccgcgcgc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCGCCGACGATG(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

ORIGIN  
 Query Match 1.2%; Score 25; DB 6; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||||  
 Db 69 CTACTAAAAATATATAAATTAGCTG 45

RESULT 50  
 TS9425/c 279 bp mRNA linear EST 09-FEB-1995  
 LOCUS yb63f04.s1 StrataGene ovary (#937217) Homo sapiens cDNA clone  
 DEFINITION IMAGE:75871.3, similar to contains Alu repetitive element, contains  
 PTRS MER28 repetitive element;; mRNA sequence.  
 ACCESSION TS9425  
 VERSION T59425.1 GI:661262  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 279)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,  
 Chisoe, S., Dietrich, N., Dubugue, T., Favello, A., Gish, W.,  
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
 Maritz, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rikkin, L.,  
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
 Trevisan, E., Underwood, K., Woldmann, P., Waterston, R., Wilson, R.,  
 and Werra, M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 JOURNAL Genome Res. 6 (9), 807-828 (1996)  
 MEDLINE 97044478

PUBMED 889549  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [est@wustl.wustl.edu](mailto:est@wustl.wustl.edu)  
 Insert Size: 346  
 Source: IMAGE Consortium, LLNL This clone is available royalty-free  
 through LLNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov))  
 for further information.  
 Insert Length: 346 Std Error: 0.00  
 Seq primer: -21m13  
 High quality sequence stop: 313.  
 Location/Qualifiers  
 1..279  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:505600"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:75871"  
 /sex="female"  
 /dev\_stage="49 year old"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Stratagene ovary (#937217)"  
 /note="Organ: ovary; Vector: Bluescript SK; Site 1: EcoRI;  
 Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt.  
 Total ovary tissue, normal, caucasian. Average insert  
 size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'  
 GAATCGGACGACG 3' -3' adaptor sequence: 5'  
 CTCGAGTTTCTTTTCTTTTCTTTT 3'"

ORIGIN  
 Query Match 1.2%; Score 25; DB 7; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||||  
 Db 171 CTACTAAAAATATATAAATTAGCTG 147

RESULT 51  
 BE043767 280 bp mRNA linear EST 08-JUN-2000  
 LOCUS hk43d02.y1 NCI CGAP Ov34 Homo sapiens cDNA clone IMAGE:2999427 5'  
 DEFINITION similar to contains Alu repetitive element;; mRNA sequence.  
 ACCESSION BE043767  
 VERSION BE043767.1 GI:8360820  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 280)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 cDNA Library Preparation: David B. Kitzman, Ph.D.  
 cDNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Putative full length read  
 The vector to vector length is 348  
 Seq primer: -40RP from Gdbco  
 High quality sequence stop: 271.  
 Location/Qualifiers  
 1..280

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2999427"
/sex="female"
/tissue_type="borderline ovarian carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Ov34"
/note="Organ: ovary; Vector: pAMP1; mRNA made from
borderline ovarian carcinoma, cDNA made by oligo-dt
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 500 bp. Primary library,
non-amplified."

```

```

Query Match      1.2%; Score 25; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2012 CTACTAAATATTAATAATTAGCTG 2036
Db      181 CTACTAAATATTAATAATTAGCTG 205

```

```

RESULT 52      TS9481      282 bp      mRNA      linear      EST 09-FEB-1995
TS9481      YB63F04.r1 StrataGene ovary (#937217) Homo sapiens cDNA clone
LOCUS      IMAGE:75871.5' similar to contains Alu repetitive element; contains
DEFINITION      MER28 repetitive element ;, mRNA sequence.

```

```

ACCESSION      TS9481

```

```

VERSION      TS9481.1 GI:661318

```

```

KEYWORDS      EST.
SOURCE      Homo sapiens (human)

```

```

ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```

```

REFERENCE      1 (bases 1 to 282)
AUTHORS      Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chapelli,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Scheilenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevasakis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.,
and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

```

```

COMMENT

```

```

JOURNAL      Washington University School of Medicine
MEDLINE      4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
PUBMED      Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 346
            High quality sequence stops: 193 Source: IMAGE Consortium, LNL This
            clone is available royalty-free through LNL; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 346 Std Error: 0.00
            Seg primer: M13RP1
            High quality sequence stop: 193.

```

```

FEATURES

```

```

source
1..282
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:505600"
/db_xref="taxon:9606"
/clone="IMAGE:75871"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"

```

```

/clone_lib="Stratagene ovary (#937217)"
/note="Organ: ovary; Vector: Bluescript SK; Site 1: EcoRI;
Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
GAATTCGGCAGCAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTTTTT 3'"

```

```

Query Match      1.2%; Score 25; DB 7; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2012 CTACTAAATATTAATAATTAGCTG 2036
Db      99 CTACTAAATATTAATAATTAGCTG 123

```

```

RESULT 53      R82328      285 bp      mRNA      linear      EST 14-JUN-1995
R82328/c      YJ18A03.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
LOCUS      IMAGE:149068.3' similar to contains Alu repetitive element; , mRNA
DEFINITION      sequence.

```

```

ACCESSION      R82328
VERSION      R82328.1 GI:861719
KEYWORDS      EST.
SOURCE      Homo sapiens (human)

```

```

ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```

```

REFERENCE      1 (bases 1 to 285)
AUTHORS      Hillier,L., Clark,N., Dubuque,T., Ellison,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The Washu-Merck EST Project
Unpublished (1995)

```

```

TITLE      The Washu-Merck EST Project
JOURNAL      Contact: Wilson RK
COMMENT      Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 949
            High quality sequence stops: 203
            Source: IMAGE Consortium, LNL
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 949 Std Error: 0.00
            Seg primer: Promega -21ml3
            High quality sequence stop: 203.

```

```

FEATURES

```

```

source
1..285
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:560903"
/db_xref="taxon:9606"
/clone="IMAGE:149068"
/sex="female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta Nb2HP"
/note="Organ: placenta; Vector: pYT73D (pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AAGTCGAGAGATTCGCGCGCAGAGATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pYT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

```

## ORIGIN

Query Match 1.2%; Score 25; DB 7; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAAAATTAATAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 DB 160 CTACTAAAAATTAATAATTAGCTG 136

## RESULT 54

BI062031 287 bp mRNA linear EST 15-JUN-2001  
 DEFINITION IL3-UT0117-130301-499-A05 UT0117 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BI062031  
 VERSION BI062031.1 GI:14469558  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costra, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202653  
 PUBMED 10737800

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/ILICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL3&c2=IL3-UT0117-130301-499-A05&c3=2001-03-13&c4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 287.  
 Location/Qualifiers

## FEATURES

1..287  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="UT0117"  
 /note="Organ: uterus tumor; Vector: puc18, Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 1.2%; Score 25; DB 4; Length 287;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAAAATTAATAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 DB 188 CTACTAAAAATTAATAATTAGCTG 212

## RESULT 55

AW770166/c 292 bp mRNA linear EST 04-MAY-2000  
 LOCUS hK59b06.x1 NCI CGAP Lym12 Homo sapiens cDNA IMAGE:3000947 3'  
 DEFINITION similar to contains Alu repetitive element; contains element MIR repetitive element; , mRNA sequence.

ACCESSION AW770166  
 VERSION AW770166.1 GI:7702205  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 292)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)

## JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Life Technologies catalog #: 11547-015  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LMNC, send email to: info@image.lim.gov  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 272.  
 Location/Qualifiers

## FEATURES

1..292  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_image="3000947"  
 /tissue\_type="lymphoma, follicular mixed small and large cell"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI-CGAP Lym12"  
 /note="Organ: lymph node; Vector: pCMV-SPORT6; Site\_1: SmaI; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"

## ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAAAATTAATAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 DB 169 CTACTAAAAATTAATAATTAGCTG 145

## RESULT 56

BI031941 293 bp mRNA linear EST 14-JUN-2001  
 DEFINITION IL5-MT0266-110401-416-g10 MT0266 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BI031941  
 VERSION BI031941.1 GI:14438571  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costra, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE  
PUBMED  
COMMENT

20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL5&cl2=IL5-MT0266-  
110401-416-910&cl3=2001-04-11&cl4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 293.

## FEATURES

source

Location/Qualifiers

1..293  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="WT0266"  
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

## ORIGIN

Query Match 1.2%; Score 25; DB 4; Length 293;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2012 CTACTAAATAATTAATAATTAGCTG 2036  
Db 196 CTACTAAATAATTAATAATTAGCTG 220

RESULT 57  
AM832979/c 294 bp mRNA linear EST 18-MAY-2000  
LOCUS AM832979 RC3-TT0005-061099-011-d11 TT0005 Homo sapiens cDNA, mRNA sequence.  
DEFINITION AM832979  
ACCESSION AM832979.1 GI:7926953  
VERSION  
KEYWORDS  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 294)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.G.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=cl2-RC3-TT0005-061  
099-011-d11&cl3=1999-10-06&cl4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 45  
High quality sequence stop: 294.  
Location/Qualifiers

## FEATURES

source

1..294  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="TT0005"  
/note="Organ: testis; Vector: puc18; Site 1: SmaI; Site 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

## ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2012 CTACTAAATAATTAATAATTAGCTG 2036  
Db 237 CTACTAAATAATTAATAATTAGCTG 213

RESULT 58  
AA719115/c 302 bp mRNA linear EST 30-DEC-1997  
LOCUS AA719115 2h3f01.81 Soares\_pineal\_gland\_N3HPG Homo sapiens cDNA clone  
DEFINITION IMAGE:41381 3' similar to contains Alu repetitive element;; mRNA  
sequence.  
ACCESSION AA719115  
VERSION AA719115.1 GI:2732214  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 302)  
Hillier,L., Allen,M., Bowles,L., Dubque,T., Gaisel,G., Jost,S.,  
Kizman,D., Kucada,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Stepien,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
WashU-NCI human EST Project  
Unpublished (1997)  
TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: es@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40mJ3 fwd. Et from Amersham  
High quality sequence stop: 294.  
Location/Qualifiers

## FEATURES

source

1..302  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="IMAG:41381"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares\_pineal\_gland\_N3HPG"  
/note="Organ: pineal gland; Vector: pV773D (Pharmacia)  
with a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
1st strand cDNA was primed with a Not I - oligo(dT) primer

[5'] TGTTACCAATCGAGTCGAGCGCCGCCGCTTTTCTTTTCTTTT  
 3)], double-stranded cDNA was size selected, ligated to  
 Eco RI adapters (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified pRT3  
 vector (Pharmacia). Library constructed by Bento Soares  
 and M. Petim Bonaldo. "

## ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036

Db 172 CTACTAAATATATAAATTAGCTG 148

## RESULT 59

LOCUS B74377 303 bp DNA linear GSS 24-OCT-1998  
 DEFINITION CIT-HSP-2028B14.TR CIT-HSP Homo sapiens genomic clone 2028B14,  
 genomic survey sequence.

## ACCESSION

B74377

## VERSION

B74377.1 GI:2770064

## KEYWORDS

GSS.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 (bases 1 to 303)  
 Adams,M.D., Rounseley,S.D., Field,C.E., Bass,S., Linher,K.,  
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
 Simon,M., and Venter,J.C.

## AUTHORS

Simon,M., and Venter,J.C.

## TITLE

Use of a random BAC End Sequence Database for Sequence-Ready Map

## JOURNAL

Unpublished (1997)

COMMENT Other GSSs: CIT-HSP-2028B14.TF

## CONTACT

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html

## FEATURES

Seq primer: M13 Reverse

## CLASS

BAC ends.

Location/Qualifiers

1..303

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="GDB:7047976"

/db\_xref="taxon:9606"

/clone="2028B14"

/sex="Male"

/cell\_type="Sperm"

/clone\_1ib="CIT-HSP"

/note="Vector: pBeloBAC11, Site\_1: HindIII, Site\_2:

HindIII"

## ORIGIN

Query Match 1.2%; Score 25; DB 8; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036

Db 188 CTACTAAATATATAAATTAGCTG 212

## RESULT 60

A0585018/c  
 LOCUS A0585018 307 bp DNA linear GSS 07-JUN-1999  
 DEFINITION RPCI-11-451H7.TV RPCI-11 Homo sapiens genomic clone RPCI-11-451H7,  
 genomic survey sequence.

## ACCESSION

A0585018

## VERSION

A0585018.1 GI:5012064

## KEYWORDS

GSS.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 (bases 1 to 307)  
 Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and  
 Venter,J.C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

## AUTHORS

Map Building

## TITLE

Unpublished (1997)

## JOURNAL

Other GSSs: RPCI-11-451H7.TV

## COMMENT

Contact: Shaying Zhao, William Niernan, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeet@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genet cs (info@resgen.com). BAC end search page:

http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..307

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="GDB:7672974"

/db\_xref="taxon:9606"

/clone="RPCI-11-451H7"

/sex="Male"

/cell\_type="Lymphocytes"

/clone\_1ib="RPCI-11"

/note="Vector: pBAC63.6, Site\_1: EcoRI, Site\_2: EcoRI;

RPCI11 Human Male BAC Library"

## ORIGIN

Query Match 1.2%; Score 25; DB 8; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036

Db 38 CTACTAAATATATAAATTAGCTG 14

## RESULT 61

AQ104891/c  
 LOCUS AQ104891 308 bp DNA linear GSS 28-AUG-1998  
 DEFINITION HS.2166.A2.A12.MR.CIT Approved Human Genomic Sperm Library D Homo  
 sapiens genomic clone Plate=2166 Col=24 Row=A, genomic survey  
 sequence.

## ACCESSION

AQ104891

## VERSION

AQ104891.1 GI:3480247

## KEYWORDS

GSS.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 (bases 1 to 308)  
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE 99300589  
PUBMED 10449764  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 2166 row: A column: 24  
Class: BAC ends  
High quality sequence stop: 308.

FEATURES  
source 1..308  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=2166 Col=24 Row=A"  
/sex="male"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: Sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

## ORIGIN

Query Match 1.2%; Score 25; DB 8; Length 308;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
|||||  
Db 212 CTTACTAAAAATATATAAATTAGCTG 188

RESULT 62 310 bp DNA linear GSS 20-APR-1999  
LOCUS AQ083275  
DEFINITION RPCI11-57A24.TK RPCI-11 Homo sapiens genomic clone RPCI-11-57A24,  
genomic survey sequence.  
ACCESSION AQ083275  
VERSION AQ083275.1 GI:3444459  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 310)  
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
Berry,K., Granger,D., Sub,E., Wible,C., de Jong,P. and Venter,J.C.  
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building  
JOURNAL Unpublished (1998)  
COMMENT Other\_GSSs: RPCI11-57A24.TJ  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Library are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/humgen/bac\_end\_search/bac\_end\_search.html  
Class: BAC ends.

FEATURES  
source 1..310  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

/db\_xref="GDB:7521527"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-57A24"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPCI11 Human Male BAC Library"

## ORIGIN

Query Match 1.2%; Score 25; DB 8; Length 310;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2012 CTTACTAAAAATATATAAATTAGCTG 2036  
|||||  
Db 201 CTTACTAAAAATATATAAATTAGCTG 225

RESULT 63 312 bp mRNA linear EST 07-FEB-2003  
LOCUS BX109463/c  
DEFINITION BX109463 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGp998L20471 ; IMAGE:233635, mRNA sequence.  
ACCESSION BX109463  
VERSION BX109463  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 312)  
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,  
Radelof,U., Schneider D. and Korn,B.  
TITLE Human Unigeneset - RZPD3  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD: IMAGp998L20471.  
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No.972)  
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heuberweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r, Primer sequence: TTTCACACAGGAACAGCTATGAC.

FEATURES  
source 1..312

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGp998L20471 ; IMAGE:233635"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
(5' AACTGAGAAATTATTAAGATCTTTTCTTTTCTTTT 3'),  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN



Query Match 1.2%; Score 25; DB 5; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||  
 85 CTACTAAAAATATATAAATTAGCTG 61

RESULT 64  
 A0029448/c 312 bp DNA linear GSS 14-APR-1999  
 LOCUS RPL11-38N10.7P RPL11-11 Homo sapiens genomic clone RPL11-38N10,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0029448  
 VERSION A0029448.1 GI:3274579  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 312)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Baes,S., Linher,K.,  
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and  
 Venter,J.C.  
 Venter,J.C.  
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)  
 JOURNAL Unpublished (1998)  
 COMMENT Other GSSs: RPL11-38N10.TV  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are derived from the human BAC library RPL11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@edj.med.buffalo.edu). Clones may be purchased from  
 BACRAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genetics (info@resgen.com). BAC end search page:  
 http://www.tigr.org/cdb/humgen/bac\_end\_search/bac\_end\_search.html  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..312  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7514529"  
 /db\_xref="taxon:9606"  
 /clone="RPL11-38N10"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /clone\_1ib="RPL11"  
 /note="Vector: PBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPL11 Human Male BAC Library"

ORIGIN  
 Query Match 1.2%; Score 25; DB 8; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||  
 118 CTACTAAAAATATATAAATTAGCTG 94

RESULT 65  
 AA625620 317 bp mRNA linear EST 02-MAR-1998  
 LOCUS ad1011.s1 Soares,NbHFB Homo sapiens cDNA clone IMAGE:877820 3',  
 DEFINITION similar to contains Alu repetitive element;; mRNA sequence.  
 ACCESSION AA625620  
 VERSION AA625620.1 GI:2538007

KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 317)  
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisler,G., Jost,S.,  
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marris,M.,  
 Martin,J., Moore,B., Schellenderg,K., Steptoe,M., Tan,F.,  
 Thesling,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
 WashU-NCI human EST Project  
 TITLE Unpublished (1997)  
 JOURNAL Contract: Wilson RK  
 COMMENT Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.jnl.gov) for further information.  
 Insert Length: 862 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amerisham  
 High quality sequence stop: 282.

FEATURES  
 source  
 Location/Qualifiers  
 1..317  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:877820"  
 /dev\_stage="15 wk post natal"  
 /lab\_host="DH10B"  
 /clone\_1ib="Soares,NbHFB"  
 /note="Organ: Whole Brain; Vector: pTT73D-Pac (Pharmacia)  
 with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 15' AACTGGAAGATTCGCGCCGCAATATTTTATTTTATTTT 3',  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pTT73 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
 Query Match 1.2%; Score 25; DB 1; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
 |||  
 170 CTACTAAAAATATATAAATTAGCTG 146

RESULT 66  
 BF836228 317 bp mRNA linear EST 13-JUN-2001  
 LOCUS QV3-HT1016-171100-474-C03 HT1016 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF836228  
 ACCESSION BF836228.1 GI:12187942  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 317)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.U., Soares,F., Brentani,R.R., Reis,I.F., de Souza,S.J. and  
 Simpson,A.U.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV3&ct=QV3-HT1016-  
 171100-474-c03&ct3=2000-11-17&ct4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 7  
 High quality sequence stop: 317.  
 Location/Qualifiers  
 1..317  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT1016"  
 /note="Organ: head,neck; Vector: puc18; Site: 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

ORIGIN  
 Query Match 1.2%; Score 25; DB 2; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 Db 260 CTACTAAATATATAAATTAGCTG 284

RESULT 67  
 AA632845 319 bp mRNA linear EST 30-OCT-1997  
 LOCUS np846f02.s1 NCI\_CGAP\_Thy1 Homo sapiens cDNA clone IMAGE:1133019  
 DEFINITION similar to contains Alu repetitive element; contains element TAR1  
 AA632845  
 VERSION AA632845.1 GI:2556259  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 319)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
 Emert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bdrp/image/image.html  
 Insert Length: 563 Std Error: 0.00  
 Seq primer: -40m13 fwd. BT from Amerisham  
 High quality sequence stop: 303.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source Location/Qualifiers  
 1..319  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:1133019"  
 /tissue\_type="thyroid"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Thy1"  
 /note="Vector: pAMP10; mRNA made from invasive thyroid  
 tumor, cDNA made by oligo-dT priming. Non-directionally  
 cloned. Size-selected on agarose gel, average insert size  
 600 bp. Reference: Krizman et al. (1996) Cancer Research  
 56:5380-5383."

ORIGIN  
 Query Match 1.2%; Score 25; DB 1; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
 ||||||||||||||||||||||||||||  
 Db 175 CTACTAAATATATAAATTAGCTG 151

RESULT 68  
 BG008292 320 bp mRNA linear EST 24-JAN-2001  
 LOCUS CM2-GN0221-291100-553-f06 GN0221 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BG008292  
 ACCESSION BG008292.1 GI:12453335  
 VERSION  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 320)  
 Nagai,M.A., de Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Golman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.V., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL.  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM2&ct=CM2-GN0221-  
 291100-553-f06&ct3=2000-11-29&ct4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 3  
 High quality sequence stop: 320.  
 Location/Qualifiers  
 1..320  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="GN0221"  
 /note="Organ: placenta\_normal; Vector: puc18; Site: 1:  
 SmaI; Site 2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

	Matches	25;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
OY		2012 <sup>3</sup>	CTACTAAATAATTTAGTCG		2036					
Dδ		154	CTACTAAATAATTTAGTCG		130					

RESULT 70	LOCUS	AM938545	AM938545	329 bp	mRNA	linear	EST 30-MAY-2001
DEFINITION	PM4-DT0061-010200-003-F12	DT0061	Homo sapiens	CDNA, mRNA	sequence.		
ACCESSION	AM938545						
VERSION	AM938545.1	GI:8113988					
FEATURES							

ORGANISM	REFERENCE
<i>Homo sapiens</i>	
Eukaryote; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;	
Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.	
1 (bases 1 to 329)	
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,	
Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,	
Costa, M. C., de Souza, M. C., de Souza, M. C., de Souza, M. C.,	

**TITLE** Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed

MEDLINE 20202663  
PUBMED 10737800

**Brazil**  
**Tel: +55-11-2704922**  
**Fax: +55-11-2707001**

```
FEATURES
  source
    1. 329
    /organism="Homo sapiens"
    /mol_type="mrna"
    /db_xref="taxon:9606"
    /dev_stage="Adult"
```

/clone\_id="D10061"  
 /note="Organ: denis; draash; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent applicator  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

LOCUS	R08502	331 bp	mRNA	linear	EST 05-APR-1995
DEFINITION	yes5f08.r1 Soares fetal liver spleen INFUS Homo sapiens CDNA clone				



Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@cnsg.sh.cn

This clone is available at CHGC in Shanghai.

## FEATURES

source

1..339

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="cdaUA12"  
/issue\_type="pneocromocytoma"  
/dev\_stage="Adult"  
/lab\_host="BM25.8"  
/clone\_1ib="cda"  
/note="Vector: pTriplex2; Site\_1: sf1a; Site\_2: sf1b"

## ORIGIN

Query Match

Best Local Similarity 1.2%; Score 25; DB 1; Length 339;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

2012 CTACTAAAAATATAAATTAGCTG 2036

Db

46 CTACTAAAAATATAAATTAGCTG 70

RESULT 74

BX480803/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 339)  
Wambuit,R., Heubner,D., Mewes,H.W., Weill,B., Amid,C., Osanger,A.,  
Fobd,G., Han,M. and Wiemann,S.  
EST (Wambuit,R., Heubner,D., Mewes,H.W., Weill,B., Amid,C., et al.)  
Unpublished (2003)  
Contract: MIPS

MIPS  
Ingelstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by AGOWA (Berlin/Germany) within the CDNA sequencing  
consortium of the German Genome Project.  
No 81 sequence available.  
This clone (DKFZ686820223) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

## FEATURES

source

1..339

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZ686820223"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/clone\_1ib="686 (synonym: h1cc3)"  
/note="Vector: pTriplex2; Site\_1: sf1a; Site\_2: sf1b;  
CDNA-collection"

## ORIGIN

Query Match

Best Local Similarity 1.2%; Score 25; DB 5; Length 339;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAAAATATAAATTAGCTG 2036

Db 193 CTACTAAAAATATAAATTAGCTG 169

RESULT 75

B83941

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Use of BAC End Sequences for Sequence-Ready Map Building (1998)  
Unpublished (1998)  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: 17  
Class: BAC ends.

## FEATURES

source

1..339

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:750777"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-21B2"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1ib="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

## ORIGIN

Query Match

Best Local Similarity 1.2%; Score 25; DB 8; Length 339;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

2012 CTACTAAAAATATAAATTAGCTG 2036

Db

118 CTACTAAAAATATAAATTAGCTG 142

RESULT 76

BE071881/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Human sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 346)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,U.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL  
MEDLINE  
PUBMED

20202663  
10737800

COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=62=RC2-BT0522-120  
200-014-e04&t3=2000-02-12&t4=1)

Seq primer: puc 18 forward  
High quality sequence start: 16  
High quality sequence stop: 346.  
Location/Qualifiers

FEATURES  
source

1.346  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BT0522"

/note="Organ: breast; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORSTS PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATATAAATTAGCTG 2036

Db 89 CTAATAAATATAAATTAGCTG 65

RESULT 77

B90305/c

DEFINITION B90305 346 bp DNA linear GSS 25-UTN-1998  
CIT-HSP-2174K6.TR CIT-HSP Homo sapiens genomic clone 2174K6,  
genomic survey sequence.

ACCESSION

B90305

VERSION

B90305.1

KEYWORDS

GSS.

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mddams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13 Reverse  
Class: BAC ends.

FEATURES  
source

location/Qualifiers  
1.346  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="GDB:7104688"  
/db\_xref="taxon:9606"  
/clone="2174K6"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

ORIGIN

Query Match 1.2%; Score 25; DB 8; Length 346;  
Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAATAAATATAAATTAGCTG 2036

Db 259 CTAATAAATATAAATTAGCTG 225

RESULT 78

BX953655

DEFINITION BX953655 347 bp mRNA linear EST 01-MAR-2004  
DKFZ78101047.r1 781 (synonym: h1ccc4) Homo sapiens cDNA clone  
DKFZ78101047.5, mRNA sequence.

ACCESSION

BX953655

VERSION

BX953655.1

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

FEATURES

source

location/Qualifiers  
1.347  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZ78101047"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="781 (synonym: h1ccc4)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb;  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No si sequence available.  
This clone (DKFZ78101047) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

CONTACT: Mark Adams  
Department of Bukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA

ORIGIN CDNA-collection"

Query Match 1.2%; Score 25; DB 5; Length 347;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 35 CTACTAAAAATATATAAATTAGCTG 59  
|||||

RESULT 79  
AA837035/c 348 bp mRNA linear EST 24-AUG-1998  
LOCUS ogo9f11.s1 NCI\_CGAP Lu6 Homo sapiens CDNA clone IMAGE:1439373  
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.  
ACCESSION AA837035  
VERSION AA837035.1 GI:2912234  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE NCI\_CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
AUTHORS 1 (bases 1 to 348)  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapb-r@mail.nih.gov  
Tissue Procurement: Michael J. Kelley, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.lnl.gov/bbrp/image/image.html  
Insert Length: 542 Std Error: 0.00  
Seq primer: -40m3 fwd. ET from Amersham  
High quality sequence stop: 328.  
Location/Qualifiers  
1. 348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1439373"  
/feature\_type="small cell carcinoma"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP Lu6"  
/note="Organ: lung; Vector: pCMV-SPORT4; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.3 kb."

ORIGIN

Query Match 1.2%; Score 25; DB 1; Length 348;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 176 CTACTAAAAATATATAAATTAGCTG 152  
|||||

RESULT 80  
BF081260 348 bp mRNA linear EST 18-OCT-2000  
LOCUS RC6-AN0063-120900-021-F04 AN0063 Homo sapiens CDNA, mRNA sequence.  
DEFINITION RC6-AN0063-120900-021-F04 AN0063 Homo sapiens CDNA, mRNA sequence.  
ACCESSION BF081260  
VERSION BF081260.1 GI:10875090  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 348)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Britones, M.R.,  
Nagai, M.A., de Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.U., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
JOURNAL MEDLINE  
PUBMED 20202663  
10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPERBP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=ct2=RC6-AN0063-120  
900-021-F04&t3=2000-09-12&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 24  
High quality sequence stop: 347.  
Location/Qualifiers  
1. 348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="AN0063"  
/note="Organ: amion normal; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAAATTAGCTG 2036  
DB 32 CTACTAAAAATATATAAATTAGCTG 56  
|||||

RESULT 81  
R66389 349 bp mRNA linear EST 30-MAY-1995  
LOCUS y13le03.s1 Soares placenta Nb2HP Homo sapiens CDNA clone  
DEFINITION IMAGE:140860 3' similar to contains Alu repetitive element; contains  
MER22 repetitive element;; mRNA sequence.  
ACCESSION R66389  
VERSION R66389.1 GI:839027  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 349)  
AUTHORS Hillier, L., Clark, N., Dubugue, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevasakis, E., Waterston, R., Willamson, A., Wohlmann, P. and

TITLE  
JOURNAL  
COMMENT  
Wilson, R.  
The Mashu-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1074  
High quality sequence stops: 210  
Source: IMAGE Consortium, LNL  
This clone is available royally-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1074 Std Error: 0.00  
Seq primer: Promega -21m13  
High quality sequence stop: 210.

FEATURES  
source  
1. .349  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:549480"  
/db\_xref="taxon:9606"  
/clone="IMAGE:140860"  
/sex="Female"  
/dev\_stage="Placenta obtained at birth (full term)"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares placenta Nb2HP"  
/note="Organ: Placenta; Vector: pT7T3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
AAGTGAAGAATTCGCGCGCGAGAAATTTTCTTTTCTTTT 3']  
, double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN  
Query Match 1.2%; Score 25; DB 7; Length 349;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 167 CTACTAAATATATAAATTAGCTG 143

RESULT 82 352 bp mRNA linear EST 05-OCT-1995  
H58252  
LOCUS yroeh04.s1 Soares fetal liver spleen INFUS Homo sapiens cDNA clone  
DEFINITION IMAGE:204535 3' similar to contains Alu repetitive element;; mRNA  
sequence.  
H58252  
VERSION H58252.1 GI:1011084  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 352)  
Hiller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.

TITLE  
JOURNAL  
COMMENT  
The Mashu-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800

Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 700  
High quality sequence stops: 297  
Source: IMAGE Consortium, LNL  
This clone is available royally-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 700 Std Error: 0.00  
Seq primer: Promega -21m13  
High quality sequence stop: 297.  
Location/Qualifiers  
1. .352  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3773666"  
/db\_xref="taxon:9606"  
/clone="IMAGE:204535"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFUS"  
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGAAGAATTAATTAAGATCTTTTCTTTTCTTTT 3']  
, double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 1.2%; Score 25; DB 7; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2012 CTACTAAATATATAAATTAGCTG 2036  
|||||  
Db 185 CTACTAAATATATAAATTAGCTG 209

RESULT 83 353 bp mRNA linear EST 20-AUG-1997  
AA225629  
LOCUS nc08a04.r1 NCI CGAP P11 Homo sapiens cDNA clone IMAGE:1007502  
DEFINITION similar to gb:K773J8\_rnal BAND 3 ANION TRANSPORT PROTEIN  
(HUMAN); contains Alu repetitive element;; mRNA sequence.  
AA225629  
VERSION AA225629.1 GI:1846955  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

REFERENCE  
AUTHORS  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 353)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

TITLE  
JOURNAL  
COMMENT  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,  
M.D., Michael Emerit-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Kitzman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[www.bio.lnl.gov/bbrp/image.html](http://www.bio.lnl.gov/bbrp/image.html)  
Insert Length: 646 Std Error: 0.00  
Seq primer: -28m13 rev1 ET from Amersham.  
Location/Qualifiers  
1. .353

FEATURES  
source



```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_xref="IMAGE:1007502"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pri"
/notes="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptor, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."

```

## ORIGIN

```

Query Match      1.2%; Score 25; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2012 CTACTAAATATTAATAATTAGCTG 2036
Db      167 CTACTAAATATTAATAATTAGCTG 143

```

```

RESULT 84
LOCUS   A0705989/c      354 bp      DNA      linear      GSS 07-UTL-1999
DEFINITION HS_3556 A2_E06_S06B RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=1132 Col=12 Row=I, genomic survey sequence.
ACCESSION A0705989
VERSION   A0705989.1 GI:5415415
KEYWORDS GSS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 354)
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pierer de Jong
(pletier@u.washington.edu). Clones may be purchased from
BACRAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end web Server:
http://www.htsc.washington.edu
Plate: 1132 row: I column: 12
Seg primer: SP6
Classes: BAC ends
High quality sequence strop: 354.
Location/Qualifiers
1..354
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

```

```

FEATURES
source
1..354
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

/clone="Plate=1132 Col=12 Row=I"
/sex="Male"
/clone_lib="RPCI-11 Human Male BAC Library"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"

```

## ORIGIN

```

Query Match      1.2%; Score 25; DB 8; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2012 CTACTAAATATTAATAATTAGCTG 2036
Db      238 CTACTAAATATTAATAATTAGCTG 274

```

```

RESULT 85
LOCUS   BF820331      357 bp      mRNA      linear      EST 13-JAN-2001
DEFINITION CM0-RT0018-181100-706-e12 RT0018 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF820331
VERSION   BF820331.1 GI:12158774
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 357)
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=CM0&t2=CM0-RT0018-
181100-706-e12&t3=2000-11-18&t4=1)
Seg primer: puc 18 forward
High quality sequence strop: 109
High quality sequence strop: 176.
Location/Qualifiers
1..357
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="RT0018"
/notes="Organ: kidney tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORSSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

## REFERENCE

```

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

```

## FEATURES

```

source
1..357
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="RT0018"
/notes="Organ: kidney tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORSSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

## ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAAGTAAATATATAAATTAGCTG 2036  
 |||||  
 291 CTAAGTAAATATATAAATTAGCTG 315

RESULT 86  
 AO374407 357 bp DNA 1linear GSS 20-MAY-1999  
 LOCUS RPCI11-157B22.TV RPCI-11 Homo sapiens genomic clone RPCI-11-157B22,  
 DEFINITION genomic survey sequence.  
 ACCESSION AO374407  
 VERSION AO374407.1 GI:4345430  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 357)  
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and  
 Venter,J.C.  
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 Unpublished (1997)  
 Other\_GSSs: RPCI11-157B22.TJ  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@edj.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genetics (info@resgen.com). BAC end search page:  
 http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 source Location/Qualifiers  
 1..357  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:755949"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-157B22"  
 /sex="Male"  
 /cell\_type="lymphocytes"  
 /clone\_lib="RPCI-11"  
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
 RPCI11 Human Male BAC Library"

ORIGIN  
 Query Match 1.2%; Score 25; DB 8; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAAGTAAATATATAAATTAGCTG 2036  
 |||||  
 216 CTAAGTAAATATATAAATTAGCTG 240

RESULT 87  
 AO507012 357 bp DNA 1linear GSS 29-APR-1999  
 LOCUS RPCI-11-314A20.TJ RPCI-11 Homo sapiens genomic clone  
 DEFINITION RPCI-11-314A20, genomic survey sequence.  
 ACCESSION AO507012  
 VERSION AO507012.1 GI:4711759

KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 357)  
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and  
 Venter,J.C.  
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 Unpublished (1997)  
 Other\_GSSs: RPCI-11-314A20.TV  
 Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@edj.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genet cs (info@resgen.com). BAC end search page:  
 http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source Location/Qualifiers  
 1..357  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7620211"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-314A20"  
 /sex="Male"  
 /cell\_type="lymphocytes"  
 /clone\_lib="RPCI-11"  
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
 RPCI11 Human Male BAC Library"

ORIGIN  
 Query Match 1.2%; Score 25; DB 8; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTAAGTAAATATATAAATTAGCTG 2036  
 |||||  
 327 CTAAGTAAATATATAAATTAGCTG 351

RESULT 88  
 BE185887/c 359 bp mRNA 1linear EST 22-JUN-2000  
 LOCUS IL5-HT0731-120500-085-e10 HT0731 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BE185887  
 ACCESSION BE185887.1 GI:8665071  
 VERSION EST.  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 359)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bais,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.D., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 JOURNAL MEDLINE 20202663  
 PUBMED 10737800

COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=6c2=IL5-HT0731-120500-085-e10&cl3=2000-05-12&cl4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 24  
High quality sequence stop: 319.  
Location/Qualifiers  
1. 359  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0731"  
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ONESTEP PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN  
Query Match 1.2%; Score 25; DB 2; Length 359;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
Db 70 CTACTAAATATATAAATTAGCTG 46

RESULT 89  
B2605649/c 361 bp DNA linear GSS 08-JUN-2003  
LOCUS B2605649 Human MCF7 breast cancer cell line library (MCF7\_1) Homo  
DEFINITION B2605649 sapiens genomic clone MCF7\_1-8b17, genomic survey sequence.  
ACCESSION B2605649  
VERSION B2605649.1 GI:31514210  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 361)  
Volik S., Zhao S., Chin K., Brubner J.H., Herndon D.R., Tao Q.,  
Gray J.W. and Collins C.  
End-sequence profiling: Sequence-based analysis of aberrant genomes  
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
22709111  
PUBMED 12788976  
COMMENT Contact: Volik SV  
Collin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7066  
Fax: 415 502 5665  
Email: svolik@cc.ucsf.edu  
This clone is available from Amplicon Express  
<http://www.genomex.com>  
Classes: BAC ends.  
Location/Qualifiers  
1. 361  
/organism="Homo sapiens"

FEATURES  
source

ORIGIN  
Query Match 1.2%; Score 25; DB 8; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAATATATAAATTAGCTG 2036  
Db 329 CTACTAAATATATAAATTAGCTG 305

RESULT 90  
A1124635 363 bp mRNA linear EST 11-SEP-1998  
LOCUS A1124635  
DEFINITION amsg02.x1 Johnston frontal cortex Homo sapiens cDNA clone  
IMAGE:1539890 3', mRNA sequence.  
ACCESSION A1124635  
VERSION A1124635.1 GI:3593149  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 363)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gaisel, G., Jost, S.,  
Kritman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F.,  
Theising, B., White, Y., Wylie, T., Waterson, R. and Wilson, R.  
Unpublished (1997)  
Washington University School of Medicine  
Contact: Wilson RK  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watsn.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: 40m3 fwd. ET from Amersham  
High quality sequence stop: 359.  
Location/Qualifiers  
1. 363  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1539890"  
/sex="male"  
/tissue\_type="pooled frontal lobe"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_lib="Johnston frontal cortex"  
/note="Organ: Brain; Vector: Bluescript SK-; Site 1:  
Scor1; Stanley Neuropathology Consortium  
([www.stanleylab.org](http://www.stanleylab.org)) brains S-58, S-65, S-67, S-78.  
Random + oligo-dT primed into Scor1 site of ZAP II Vector.  
Mass excised. Avg insert length 1.9kb. Custom library  
provided by Dr. Nancy Johnston [(410) 614-3918,  
[nj@welchlink.welch.jhu.edu](mailto:nj@welchlink.welch.jhu.edu)]."

ORIGIN  
Query Match 1.2%; Score 25; DB 1; Length 363;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAATATATAAATTAGCTG 2036  
 Db 333 CTACTAAATATATAAATTAGCTG 357

RESULT 91  
 LOCUS AI283938/c  
 DEFINITION AT283938 366 bp mRNA linear EST 21-DEC-1998  
 qt71e11.x1 NCI\_CGAP\_Eso2 Homo sapiens cDNA clone IMAGE:1960748.3'  
 similar to contains Alu repetitive element; contains element PTRS  
 repetitive element ; mRNA sequence.

ACCESSION AI283938  
 VERSION AI283938  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 366)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip  
 Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 495 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 335.

FEATURES  
 source  
 Location/Qualifiers  
 1..366  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1960748"  
 /cissue\_type="squamous cell carcinoma"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Eso2"  
 /note="Organ: esophagus; Vector: pCMV-SPORT6, Site 1:  
 SalI; Site 2: NotI; Cloned unidirectionally. Primer:  
 oligo dt. Average insert size 1.1 kb. Life Technologies  
 catalog #: 11502-010"

ORIGIN  
 Query Match 1.2%; Score 25; DB 1; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAATATATAAATTAGCTG 2036  
 Db 168 CTACTAAATATATAAATTAGCTG 144

RESULT 92  
 LOCUS BM671780/c  
 DEFINITION UI-E-CQ1-aggd-p-07-0-UI-81 UI-E-CQ1 Homo sapiens cDNA clone  
 BM671780  
 VERSION BM671780  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 368)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 20-299, >Alu (matched complement)  
 Seq primer: M13 Forward  
 POLYA=Yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..368  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-CQ1-aggd-p-07-0-UI"  
 /cissue\_type="optic nerve"  
 /dev stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-CQ1"  
 /note="Organ: eye; Vector: p773-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-CQ1 is a normalized cDNA library containing the  
 following tissue(s): optic nerve. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into p773-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (drr)18 tail. The sequence tag for this library is  
 CCATTAGTG. This library was created for the program, Gene  
 Discovery in the Visual System, supported by National Eye  
 Institute (NEI).  
 TAG TISSUE=human optic nerve  
 TAG\_LIB=UI-E-CQ1  
 TAG\_SEQ=CCATTAGTG"

ORIGIN  
 Query Match 1.2%; Score 25; DB 4; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2012 CTACTAAATATATAAATTAGCTG 2036  
 Db 184 CTACTAAATATATAAATTAGCTG 160

RESULT 93  
 LOCUS R08675/c  
 DEFINITION R08675 369 bp mRNA linear EST 05-APR-1995  
 YF20908.r1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone  
 IMAGE:127454.5' similar to contains Alu repetitive element;; mRNA  
 sequence.

ACCESSION R08675  
 VERSION R08675  
 KEYWORDS EST.  
 SOURCE Homo sapiens

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 369)

**AUTHORS** Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawking, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

**TITLE** The Mashu-Merck EST Project

**JOURNAL** Unpublished (1995)

**COMMENT** Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

**Insert Size:** 1324

**High quality sequence stops:** 327 Source: IMAGE Consortium, LNL. This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

**Insert Length:** 1324 Std Error: 0.00

**Seq primer:** M13RP1

**High quality sequence stop:** 327.

**Location/Qualifiers**

**FEATURES**

**source**

1..369

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:479615"

/db\_xref="taxon:9606"

/clone="IMAGE:127454"

/sex="male"

/dev\_stage="20 week-post conception fetus"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: pRT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pRT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

**ORIGIN**

**Query Match** 1.2%; Score 25; DB 7; Length 369;

**Best Local Similarity** 100.0%; Pred. No. 1.1;

**Matches** 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 2012 CTACTAAATATTAATAATTAGCTG 2036

|||||

67 CTACTAAATATTAATAATTAGCTG 43

**RESULT 94**

AV646212 371 bp mRNA linear EST 15-JAN-2002

**LOCUS** AV646212 GUC Homo sapiens cDNA clone GLCAKPO5 3', mRNA sequence.

**DEFINITION** AV646212

**ACCESSION** AV646212

**VERSION** AV646212.1 GI:9867226

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 371)

**AUTHORS** Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Qu, J., Liu, F., Huang, Q., Chen, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, J., Chen, Z., and Han, Z.

**Insight into hepatocellular carcinogenesis at transcriptome level**

by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

**JOURNAL** MEDLINE 21625106

**PUBMED** 11752456

**COMMENT** Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn

**High quality sequence stop:** 201.

**Location/Qualifiers**

**FEATURES**

**source**

1..371

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="GLCAKPO5"

/tissue\_type="corresponding non cancerous liver tissue"

/dev\_stage="Adult"

/lab\_host="SOLR"

/clone\_lib="GLC"

/note="Vector: pBluescript sk(-); Site 1: EcoRI; Site 2: XhoI"

**ORIGIN**

**Query Match** 1.2%; Score 25; DB 1; Length 371;

**Best Local Similarity** 100.0%; Pred. No. 1.1;

**Matches** 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 2012 CTACTAAATATTAATAATTAGCTG 2036

|||||

**DB** 129 CTACTAAATATTAATAATTAGCTG 105

**RESULT 95**

AA195886 372 bp mRNA linear EST 22-JAN-1997

**LOCUS** AA195886/c

**DEFINITION** zp92g10.r1 Stratagene Hela cell s3 937226 Homo sapiens cDNA clone IMAGE:627714 5' similar to contains Alu repetitive element; contains element PIR5 repetitive element; mRNA sequence.

**ACCESSION** AA195886

**VERSION** AA195886.1 GI:1791460

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 372)

**AUTHORS** Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

**Generation and analysis of 280,000 human expressed sequence tags**

**TITLE** EST.

**JOURNAL** Genome Res. 6 (9), 807-828 (1996)

**MEDLINE** 97044478

**PUBMED** 889549

**COMMENT** Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

**High quality sequence stop:** 201.

**Location/Qualifiers**

**FEATURES**

**source**

1..372

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5046412"
/db_xref="taxon:9606"
/clone="IMAGE:627714"
/sex="female"
/dev_stage="HeLa S3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene HeLa cell s3 937216"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. HeLa S3
epithelioid carcinoma cells grown to semi-confluency
without induction. Average insert size: 1.5 kb; Uni-ZAP XR
Vector. ~5' adaptor sequence: 5' GAATTCGGCAGCG 3' ~3'
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

```

## ORIGIN

```

Query Match      1.2%; Score 25; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAATATATAAATTAGCTG 2036
        |||||
        178 CTACTAAATATATAAATTAGCTG 154

```

```

RESULT 96
LOCUS      BE090515      374 bp      mRNA      linear      EST 12-JUN-2000
DEFINITION RC6-BT0717-110400-011-H09 BT0717 Homo sapiens CDNA, mRNA sequence.
ACCESSION  BE090515
VERSION     BE090515.1 GI:8480966
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 374)
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,
Nagai M.A., da Silva W. Jr., Zago M.A., Bordin S., Coetra F.F.,
Goldman G.H., Carvalho A.F., Matsukuma A., Bala G.S., Simpson D.H.,
Brunstein A., de Oliveira P.S., Bucher P., Jongeneel C.V.,
O'Hare M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and
Simpson A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

```

TITLE Sequence sequencing of the human transcriptome with ORF expressed

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT

CONTACT: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=4&t=RC6-BT0717-110  
400-011-H09&t3=2000-04-11&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 36  
High quality sequence stop: 373.

FEATURES

source

```

1..374
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0717"
/notes="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived

```

from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

```

Query Match      1.2%; Score 25; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2012 CTACTAAATATATAAATTAGCTG 2036
        |||||
        209 CTACTAAATATATAAATTAGCTG 233

```

```

RESULT 97
LOCUS      AA640617/c      376 bp      mRNA      linear      EST 27-OCT-1997
DEFINITION nr22a01.r1 NCI CGAP Pr2 Homo sapiens CDNA clone IMAGE:1168680
similar to contains 'Alu repetitive element; contains element MER22
repetitive element'; mRNA sequence.
ACCESSION  AA640617
VERSION     AA640617.1 GI:2565867
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 376)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: W. Maxton Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNM at:
www.bio.lnl.gov/bbrp/image/image.html
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 260.
Location/Qualifiers
1..376
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1168680"
/sex="Male"
/dev_stage="45 years old"
/lab_host="NCI CGAP Pr2"
/notes="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected preneoplastic cells
histologically-determined to be prostatic intraepithelial
neoplasia 2 (PIN2) cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."

```

## ORIGIN

```

Query Match      1.2%; Score 25; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2012 CTACTAAAAATATATAATTAGCTG 2036  
 DB 171 CTACTAAAAATATATAATTAGCTG 147

RESULT 98  
 AM868627 376 bp mRNA linear EST 23-MAY-2000  
 LOCUS R01-OT0086-210300-021-d03 OT0086 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION  
 AM868627.1 GI:8048944  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 376)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 JOURNAL  
 MEDLINE  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ct2=R01-OT0086-210  
 300-021-d03&ct3=2000-03-21&ct4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 345.  
 High quality sequence stop: 345.  
 Location/Qualifiers  
 1. 376  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="OT0086"  
 /note="Organ: ovary; Vector: puc18; Site: 1: SmaI; Site: 2:  
 SmaI; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No.  
 196,716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAATTAGCTG 2036  
 DB 140 CTACTAAAAATATATAATTAGCTG 164

RESULT 99  
 AQ106283 378 bp DNA linear GSS 28-AUG-1998  
 LOCUS AQ106283/c  
 DEFINITION HS\_3051\_B1\_H11\_MF CIT Approved Human Genomic Sperm Library D Homo

ACCESSION  
 VERSION AQ106283.1 GI:3481639  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 378)  
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
 Hood, L.  
 Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 JOURNAL  
 MEDLINE  
 PUBMED 99380589  
 COMMENT Contact: Mahairas G.G., Wallace J.C., Hood L.  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Sequence Tagged Connector  
 Plate: 3051 Row: P Column: 21  
 Class: BAC ends  
 High quality sequence stop: 378.  
 Location/Qualifiers  
 1. 378  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=3051 Col=21 Row=P"  
 /sex="male"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: paeloBAC11; BAC Clones in  
 E-Coli DH10B"

ORIGIN

Query Match 1.2%; Score 25; DB 8; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2012 CTACTAAAAATATATAATTAGCTG 2036  
 DB 156 CTACTAAAAATATATAATTAGCTG 132

RESULT 100  
 AA586866/c 379 bp. mRNA linear EST 26-SEP-1997  
 LOCUS AA586866  
 DEFINITION nm67h05.s1 NCI CGAP Lari Homo sapiens cDNA clone IMAGE:1088985 3'  
 similar to contains Alu repetitive element; mRNA sequence.  
 ACCESSION AA586866  
 VERSION AA586866.1 GI:2397680  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 379)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: Stratagene, Inc.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.

